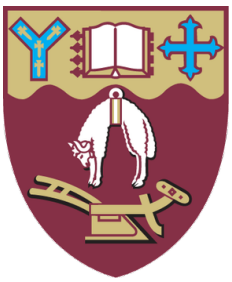
Modelling Cooperative and Competitive Behaviour in the Exploitation of Marine Resources

Michael Philip Hackney

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Mathematics



School of Mathematics and Statistics

University of Canterbury

New Zealand

July 2019

Acknowledgements

Many people have valuably contributed to this thesis and assisted me through the years it has taken me to complete it, and you all deserve tremendous thanks.

Firstly thank you to my parents Alan and Lindsey for their inexhaustible support, love, and advice throughout this experience, as have my siblings and their partners, Jonathan and Leah, and Sarah and Brodie. Your encouragement has made this all possible. Many thanks also to Bruce and Jocelyn Harray, and Keith and Merle Hackney who have always shown their support and encouragement of my academic pursuits. I hope I make you all proud.

Thank you to my girlfriend Ana. Your constant love, kindness, and understanding have been essential to me. I couldn't have finished this without you.

Thank you to all my flatmates throughout these last few years, for your friendship, understanding, and distraction. Thank you for the many great memories, I will fondly remember this part of my life.

My supervision team Mike Plank and Alex James, as well as Mike Steele, deserve thanks for their advice, inspiration, and patience throughout this project. Without your guidance through the difficulties of research, and support when I had complications, this would not have been achievable.

Thanks to my fellow students, both in the School of Maths and Stats, as well as those who I met at conferences, for their advice, interest and commiseration.

Thank you again to Jono for reading through this thesis and providing feedback and encouragement.

Further thanks to everyone at Te Pūnaha Matatini for the discussions, opportunity and guidance they have provided. I would particularly like to thank Suzi for her care and her expertise,as well as Tava for her advice on my project.

Thanks also to the team at the Social Investment Agency. My internship with you provided an invaluable breather from the academic life, and allowed me to adjust my perspective on my own goals and achievements.

This thesis was jointly funded by Te Pūnaha Matatini and the University of Canterbury.

"So long and thanks for all the fish" - Douglas Adams

Contents

Abstract

Fish are an extremely important resource worldwide, for nutritional and economic reasons. However, many fish stocks are being exploited at an ecologically unsustainable level. It is thus of importance that fisheries' management strategies can be found which will allow significant amounts of fishing to occur, while protecting these renewable resources for the future. In this thesis we are particularly concerned with situations where individual economic interests and environmental concerns coincide. We use behavioral models of individual fishers to investigate how individual profit motivation leads to aggregate fishing behavior. These are coupled with a variety of fish population models, in order to obtain a picture of how this exploitation impacts the fishery. We pay consideration to the economic value of the fishery, as well as using population abundance as a proxy for ecological health. This approach is viewed through the lens of game theory.

In Chapter 2 we consider a dynamic size spectrum PDE model of a single fish population, which is supported by a producer spectrum of plankton on which the juvenile fish prey. This size spectrum model gives the abundance of the fish population by the size of its constituent members, by tracking the transfer of biomass that occurs throughout the population due to predation, mortality, and reproduction. This size spectrum model is paired with an individual based model of many fishers exploiting the population in a small scale open access fishery. We allow individual agents to change their own size-selectivity behavior, and to make the choice of whether or not to fish in order to meet profit expectations. We find that the aggregate size-selective harvesting behavior reaches a Nash equilibrium, in which

we also observe balanced harvesting of the fish stock. Furthermore, the number of active agents in the fishery converges over time. Results from this chapter have been submitted for publication.

In Chapter 3, we consider the classic Gordon-Schaefer bioeconomic model of a fishery from a non-cooperative game theory perspective. We frame exploitation as a symmetric 2-player game in which fishers take action by selecting a harvesting intensity. For a fish population at equilibrium, we find a level of fishing effort that strictly dominates all other actions, as well as a Pareto optimal frontier where the total exploitation is equivalent to that of a monopolist. Consequently, this game is structurally equivalent to an Iterated Prisoners Dilemma. We extend our analysis to non-equilibrium populations using numerical simulations, and evaluate the relative performance of well-known strategies for an IPD (such as tit-for-tat) in these conditions.

In Chapter 4, we use a Markov decision process framework to find an optimal exploitation policy for a monopolist in a noisy environment. Optimal policies map from the population biomass to the level of fishing effort which will maximise the current and future value of the fish stock. Fish populations were modelled using a Beverton-Holt process to allow for the inclusion of noise in the stock recruitment relationship. Increasing stochasticity in the population was found to reduce the optimal fishing intensity with respect to biomass. This analysis was then extended in Chapter 5 to a Markov game situation in which there are two independent fishers acting to maximize profit. A combination of fishing policies that was a Nash equilibrium was obtained.

The thesis concludes with Chapter 6, where we summarise the thesis and present some directions for future research.

Chapter 1

Introduction

1.1 Fisheries background and overall motivation

The science of fisheries management has come a long way since T.H. Huxley claimed that "probably all the great sea fisheries, are inexhaustible: that is to say that nothing we do seriously affects the number of fish" (Huxley, 1883). Indeed, it is now widely known that fishing not only has significant and long-term effects on fish stocks, but that many fish populations are in serious decline (Myers and Worm, 2003).

Fish are an important source of nutrition, particularly in densely populated parts of the world where protein consumption levels can be low. According to the Food and Agriculture Organisation of the United Nations, 150g of fish can supply an adult with 50-60% of their daily protein requirements (FAO, 2014). Additionally, fishing is an important source of employment globally, with around 58.3 million people involved in the primary fishing sector worldwide, and with employment in the fishing sector growing faster than the global population (FAO, 2014). Overfished stocks are those harvested at ecologically unsustainable levels, providing lower yields than what are potentially attainable, and have abundances smaller than that which produces the maximum sustainable yield (MSY). 24% of global capture fisheries production comes from the ten species with the highest productivity, with most of

the stocks of these species either fully fished at the MSY level, or overfished (FAO, 2014). By rebuilding fish stocks, overall production could be increased by 16.5 million tonnes (FAO, 2014). Consequently, protecting fish stock levels while maintaining or increasing yield (Beddington et al., 2007), particularly for highly productive species, is essential to sustain the fishing sector, both as a key source of employment in developing countries, and as a provider of essential nutrition for a large number of people (Kolding et al., 2016; Mills et al., 2011).

Conventionally used top-down fisheries management strategies aim to fulfil the fundamental fisheries' goals of maximising the sustainable yield (Botsford et al., 1997), while protecting small non-spawning juvenile fish, in order to give all fish a chance to reproduce at least once (Garcia et al., 2012). These management strategies are motivated by equilibrium bases stock estimates, and can include: input controls, such as restricted fishing licenses or fixed fishing seasons; gear restrictions, such as minimum mesh sizes; and catch regulations, such as a total allowable catch (TAC) (Allison and Ellis, 2001). However, these strategies can have unintended consequences on ecosystems, including: disruption of size structure (Zhou et al., 2010; Garcia et al., 2012) in which population structure changes significantly from the unexploited state and over time the number of large fish decreases; and fisheries induced evolution (Law et al., 2013; Borrell, 2013) through the introduction of new human-induced selection pressures. Recently, the alternative approach of balancing the total fishing pressure in proportion to the natural productivity of the target species has been suggested (Garcia et al., 2012). It is hypothesized by some researchers that balanced harvesting will increase resilience of fish stocks, lower disruption, and increase yields (Jacobsen et al., 2014). There is some controversy about this hypothesis, with questions about the requirement to target juveniles, as well as the need to harvest species which are not economically viable (Burgess et al., 2015), and a lack of empirical evidence (Froese et al., 2015). Further investigation will therefore be required to establish the actual ecological efficacy of balanced harvesting, and to evaluate its economic viability in comparison to existing fisheries management strategies.

Size-based fishing regulations can be complicated, and can involve such restrictions as minimum capture size requirements, quotas, and gear restrictions, each of which can vary depending on species. These methods of fishing cause fishers to target large, mature fish, which provide a large yield in biomass terms. Big, old, fat, fertile, female fish (BOFFFFs) are mature fish that are extremely successful at reproducing (Birkeland and Dayton, 2005), and which are important for population replenishment. An outcome of selectively targeting large fish is that these mega-spawners are caught, and are thus unable to continue contributing to the population, which can lead to serious impacts on population biomass (Birkeland and Dayton, 2005). Another consequence of selectively targeting large fish of commercially valuable species is that bycatch of unwanted species and sizes is often discarded (Davis and Ryer, 2003). The proportion of bycatch to landed catch (the yield which is taken ashore) can be very high in some fisheries (Davies et al., 2009), and the mortality rate for discarded bycatch can be as high as 100% for some species (Davis and Ryer, 2003).

In addition to high-level fisheries management policies, there are a variety of fishing methods and behaviours that individual fishers are able to use. The selectivity of fishing methods refers to the targeting of specific sizes and species for harvest (MacLennan, 1992) Trawling is the use of moving boats dragging nets in order to harvest fish (Wolff et al., 2015). Sufficiently small fish are able to swim through the net-mesh, and are thus for the most part protected from being caught by trawl fishing. However, some unwanted by-catch of small fish can occur. Fish too large to fit through the mesh size will be caught, and so trawling generally selects fish larger than a certain size, determined by the net mesh size (Wolff et al., 2015). In order to combat this, additional equipment such as turtle excluder devices (TEDs) (Crowder et al., 1994) can be used. These devices are escape hatches for large organisms including loggerhead sea turtles (Crowder et al., 1994), sharks and rays, (Brewer et al., 2006) which can be fitted onto existing trawl nets in order to reduce selectivity for large fish. Gillnet fishing, set netting, and longline fishing are alternative targeting strategies in which a specific small range of sizes of fish are expected to be harvested (Wolff et al., 2015). Like trawling,

the minimum size is determined by the fineness of the net mesh. As the net does not move in a gillnet fishing strategy, large fish bounce off the net without getting trapped (Wolff et al., 2015). Particular strategies that harvesters choose to use can be investigated using individual based models. Such models can be used to simulate the behaviour of individual agents by considering the ways they interact with each other and with their environment, in order to develop a system level picture of the total impact of the exploitation of the marine species.

1.1.1 Mathematical modelling of fish populations

Because of the difficulties of observing fish species and their habitats, gaining accurate measurements of stock levels can be difficult. Historically, population estimates could be obtained through extrapolation from landed catch data or from first-hand descriptions of marine fauna (Fortibuoni et al., 2017). Data from stock assessments and surveys, and mathematical models have been used to improve stock level estimates. Mathematical models are therefore an important tool for extending or understanding of fish population dynamics (Brauer et al., 2001; Kot, 2001). Biomass in a fish population can increase through individual growth, where fish become bigger, and population growth, in which the number of fish changes.

The first significant attempt to model the growth of a fish population was by Malthus (1798), who hypothesized an exponential density-independent growth scenario, where the rate at which the population grows is proportional to the population size. An important advance in population modelling was made by Verhulst (1838) and popularised by Pearl and Reed (1920). They employed Malthus' "first principle" of population modelling to a population in which growth has an upper bound due to resource limitations, such as prey availability, habitat size, or intra-specific competition. The model imposes a carrying capacity K or maximum population size that the ecological niche can support in the long term, and means that growth is density dependent. The Verhulst-Pearl model, also called the logistic growth model, is commonly used in ecology to model populations (Kot, 2001). The

population growth at biomass u with a fixed growth rate r is

$$\frac{du}{dt} = ru\left(1 - \frac{u}{K}\right). \tag{1.1}$$

In addition to continuous time models of population dynamics, discrete time models of population growth have been used, in which organisms are grouped by generation or cohort. In discrete time models, population growth is often modelled via a "stock-recruitment" relationship, denoting how current spawning-stock biomass (used as a proxy for reproductive potential) affects the number of recruits (or surviving juveniles entering the fishery) in the following year. Key features of a stock-recruitment relationship are density-independent mortality at low stock sizes, with density-dependent mortality being able to occur over the range of stock sizes. This will result in decreasing per-capita recruitment as the current stock level becomes large. The Ricker model, which was first suggested in the context of the stock recruitment occurring in a fishery, relates the expected total number u_{n+1} of individuals in the population at a certain time, to the number of individuals u_n at a point δt earlier in time. This type of model can also be applied in an age-structured model to the size of a single cohort (single year class), as a function of the size of the previous cohort (Ricker, 1954). Like the logistic growth model, the Ricker model accounts for the intrinsic population growth rate r and the carrying capacity K:

$$u_{n+1} = u_n e^{r\left(1-\frac{u_n}{K}\right)}. \tag{1.2}$$

A similar discrete time, stock-recruitment model of successive population levels at time t apart is the Beverton-Holt model (Beverton and Holt, 2012). Both of these discrete time models are commonly used, and differ in their assumptions about how density-dependent mortality affects the juvenile fish population. The Beverton-Holt map is

$$u_{n+1} = \frac{Ku_n}{u_n + (K - u_n)\, e^{-rt}}. \tag{1.3}$$

The Beverton-Holt model assumes that juvenile mortality is linearly dependent on the current number of fish that remain alive in the cohort, and that there is always predation occurring. The Ricker model is instead based on the assumption that the juvenile mortality rate is proportional to the initial cohort size. The Beverton-Holt model can also be understood as a discrete-time equivalent of the logistic growth model. Specifically, the Beverton-Holt mapping is the fixed-time solution

$$u(t) = \frac{Ku(0)}{u(0) + (K - u(0)))e^{-rt}} \tag{1.4}$$

to the logistic growth ODE given in Equation (1.1).

The population models referred to above are unstructured models, meaning that they treat all individual population members as being identical, and only track the total population size u. Continuous time unstructured models can be written in the general form

$$\frac{du}{dt} = f(u) \tag{1.5}$$

for some function f of the current population biomass. In reality, individual fish can have different life histories, and be in different stages of their life cycles. Age-structured and size-structured models break the population down into separate categories of different age or size, and keep track of the size of each category. Age-structured population models are commonly used, with some well known examples being Leslie matrices, and the Lotka integral equation (Kot, 2001). An alternative, and potentially more useful approach is to model the fish population based on its size-structure. This class of models look at the number of individual fish of a given size (in terms of length or weight) at any given time. These size classes can be either discrete or continuous. While fish size does not directly correlate to age (Pawson, 1990), it is more directly measurable from landed catch. Of greater is interest is that the feeding preferences of individual fish are strongly based on the relative size between predator and prey (Ursin, 1973).

Simple size-structured models lump together many members of the species into a small number of groups. A simple attempt would be to solely categorise and enumerate all individuals as small fish and large fish, or as "Juvenile" and "Mature" members of the population, which could also be extended to include newborn "Recruits". More complicated size-structured models allow a continuous size range of the modelled fish population, with the population density $u(x,t)$ given at each size x at time t. A McKendrick-von Foerster partial differential equation is an example of this kind of continuous size, continuous time model (Kot, 2001):

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}(gu) - \mu u. \tag{1.6}$$

In this first order PDE, $g(x,t)$ and $\mu(x,t)$ are defined functions representing respectively the mass-specific growth rate for fish of size x at time t, and the natural rate of mortality. Size-structured models in which population dynamic processes are density dependent are called size-spectra models (Andersen et al., 2015). As the growth of individual fish is dependent on their energy intake and usage (Jobling, 1995), in a size-spectrum model the feeding preferences of fish of each size can be used to derive the growth behaviour g of the population as a whole (Law et al., 2012). The preference of larger fish to feed on individuals of a specific size can similarly be used to determine the natural mortality μ of fish that occurs due to predation. The life-history behaviours of individual fish of specific sizes are therefore used to inform population-level dynamics in a size-spectrum model. A size-spectrum version of the McKendrick-von Foerster equation can be developed by allowing population level behaviours such as growth and mortality to also depend on population density. In this case, we will have $g(u,x,t)$ and $\mu(u,x,t)$ as the mass-specific growth rate and the mortality rate, respectively.

1.1.2 Harvesting models and the economics of fisheries

Fish population models can be extended to allow for the exploitation of the fish resource for economic gain. The yield or catch is the total biomass or value of fish removed from the fish-

ery. An extremely important concept in fisheries management is the maximum sustainable yield (MSY) (Hjort, 1933). MSY is the largest yield in biomass terms that can be continually taken from the fish stock over time. Harvesting at the MSY level is consistent with keeping the population biomass at the level at where the population growth rate is at its maximum (Tsikliras and Froese, 2019). In a logistic growth model, this happens at the point where the population is at half of its carrying capacity. As the name suggests, MSY provides the largest possible sustainable yield, and as such maximises the potential usage of the fisheries resource as food. Baranov (1918) developed an early and influential harvesting model for the number of fish caught in terms of the initial population size, natural mortality, and fishing mortality, under the assumption that both sources of mortality occur simultaneously. In this model fishing is therefore not simply an additional reduction of the population size on top of natural death.

If the economic values and costs of this exploitation are taken into account, then an economic model of a fishery is obtained. The first important attempt to prioritise the economic utilisation of natural resources (including fisheries) was made by Gordon, who posited that the widely acknowledged problems of overexploitation and depletion of fish stocks are a result of an inefficient allocation of fishing efforts, caused by the lack of economic rent yielded by fisheries (Gordon, 1954). Previous fisheries management strategies aimed to obtain the largest sustainable catch without considering the human use of resources, meaning that the MSY was targeted rather than the maximum economic yield (MEY). Gordon evaluated the optimal fishing behaviour to obtain MEY as a function of fishing effort, taking into account the costs of fishing, as well as the revenues obtained from it.

The economic outcomes of fisheries are dependent on the type of access that fishers have. A single entity with the exclusive access or ability to harvest a fish population is called a monopolist. A monopolist can choose the extent to which they exploit the fish population, and so can fish in such a way to maximise their own profit from fishing. If there are two

distinct fishers (or, for example, coalitions of fishers), the fishery is a duopoly. In such a situation, each entity will wish to maximise their own returns from fishing, but must take into account the fact that their competitor can also exploit the same fish resource. This competition between fishers can result in depleted fish populations, and lower profits than with monopolist access. An open-access fishery is an unregulated situation in which there are many individual fishers with equal ability and opportunity to fish. In an open-access fishery, there are no restrictions for fishers to enter or exit the fishery, and no individual fisher can influence the market price of fish. Furthermore, individual fishers have no enforceable property rights to the fish population; if they choose not to harvest, or are unable to, there is nothing stopping another fisher from taking their "share" of the fish stock.

The harvesting of fish can be incorporated into existing population models as a harvesting term representing an external source of fish mortality. For unstructured models like those described above, the biomass harvested from the population is calculated by considering the rate F of total fishing effort. The growth of a population that is being exploited is then:

$$\frac{du}{dt} = f(u) - Fu \tag{1.7}$$

The rate at which biomass is extracted from the population is called the yield Y, and is:

$$Y = Fu \tag{1.8}$$

which means that as the fishing effort F is increased, the yield obtained from the fishery will also increase. Harvesting can also be included in size-structured population models via an additional size-dependent mortality term $F(x,t)$. That is, in such models the fishing mortality can be a function of both time and the size of fish being harvested. For example, the size-spectrum McKendrick von Foerster model with fishing will have dynamics

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(gu\right) - (\mu + F)u \tag{1.9}$$

In this structured population model, the yield is obtained by integrating over all possible body sizes:

$$Y(t) = \int_{x=0}^{\infty} F(x,t)u(x,t)dx \tag{1.10}$$

When harvesting is added to the unstructured logistic growth model, we obtain the Gordon-Schaefer (Gordon, 1954; Schaefer, 1954) model of a fishery:

$$\frac{du}{dt} = ru\left(1 - \frac{u}{K}\right) - Fu \tag{1.11}$$

This model is governed by a bioeconomic equation where the profit P obtained during a fishing period of length T when harvesting with constant effort F is

$$P = B\int_0^T Y - cFdt \tag{1.12}$$

In this model B is the market price of fish per unit biomass, which is assumed to be constant regardless of the total amount of fish caught. The constant c is the fixed cost per unit effort of fishing, and includes opportunity costs. Opportunity costs are the costs associated with putting effort towards fishing rather than other potentially beneficial alternatives. For example, a fisher could spend their time in another source of employment, and be paid for this. In such a case, the opportunity cost includes the income lost to the fisher due to the choice to fish rather than the maximally beneficial alternative. This means that if $P \geq 0$ from fishing, the fisher is receiving at least as much return from fishing than if they had chosen their best possible alternative; in such a situation a fisher will willingly continue fishing as at this level they are receiving the minimum profit required for fishing to be economically worthwhile. The MEY, as referred to earlier, is the catch level which provides the greatest difference between the total revenue obtained from fishing, and the total cost incurred. This level, as the profit maximising yield, is the monopolist's economic optimum (Gordon, 1954).

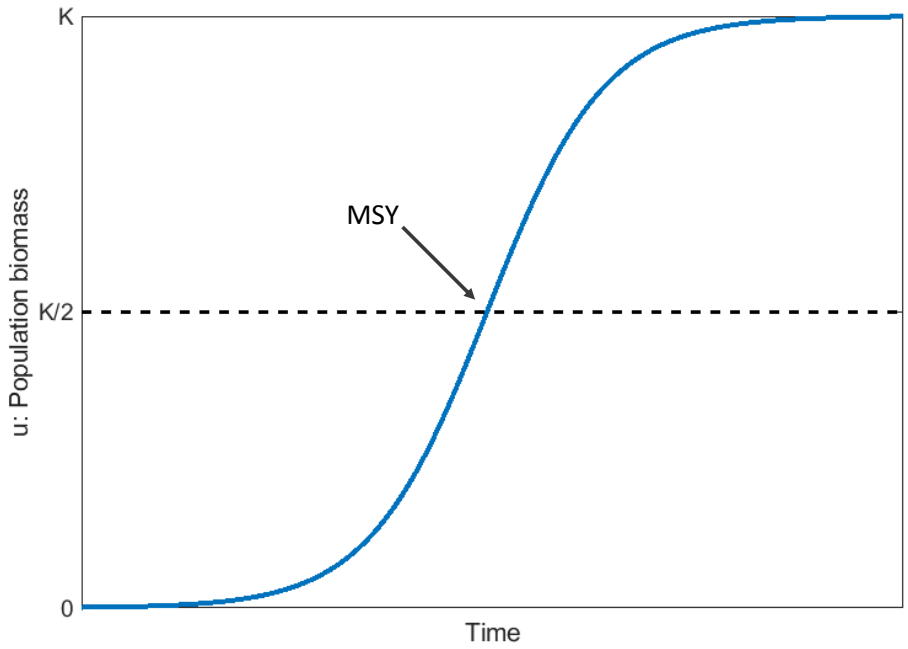


Figure 1.1: Population growth over time for a fish population with logistic growth dynamics where K is the carrying capacity of the population. The slope of the sigmoid curve gives the population growth rate; it is at its steepest when the population is at half of its carrying capacity.

The fishing effort which provides the MEY is

$$F_{MEY} = \max_{F}\{B\int_{0}^{T} Fu(t) - cFdt\}. \tag{1.13}$$

The MSY is obtained when the caught biomass is maximised, without regard to costs. This is equivalent to maximising the revenue (and not profit) obtained from fishing. This means that while we expect the yield obtained at MSY to be higher than an MEY, the profit at MSY will be lower than at MEY. The exception to this is when there are no costs associated with fishing. If $c = 0$, then revenue and profit are equal and so MSY and MEY are the same for that fishery.

In Figure 1.1, we observe the logistic growth of an unexploited fish population. The population growth rate is at its maximum when the population biomass is half of the carrying

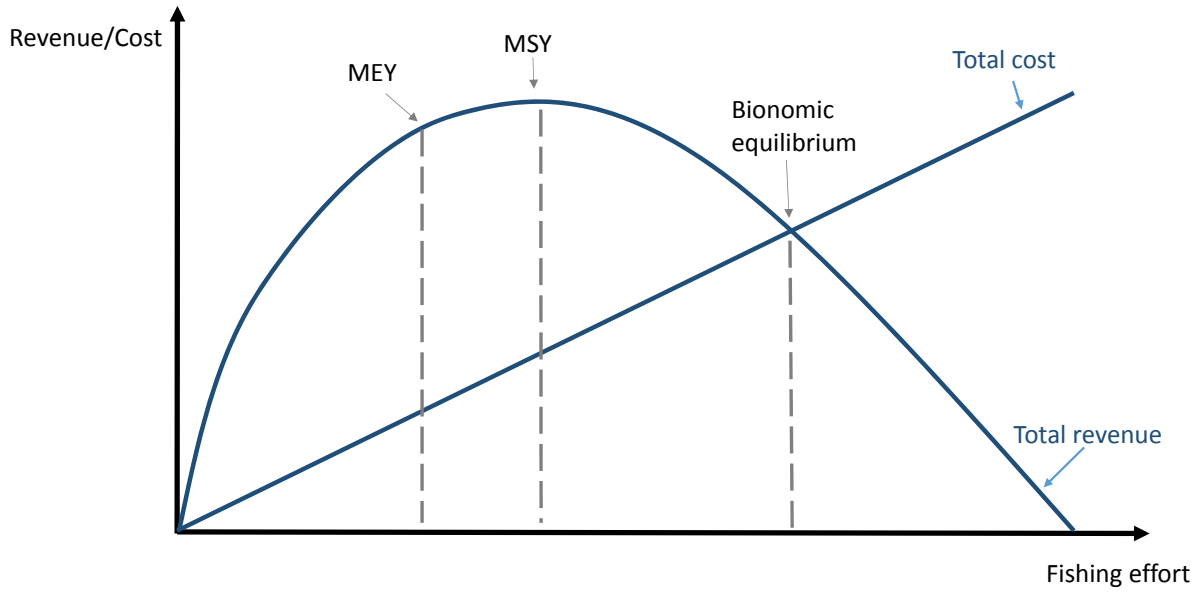


Figure 1.2: A schematic diagram of the relationship between cost and revenue in the Gordon-Schaefer bioeconomic model, as functions of fishing effort. Solutions representing MEY, MSY, and the open-access equilibrium are shown.

capacity. For a sustained yield, there will be a constant rate at which biomass is removed from the population. As such, the MSY will result when the fishing effort causes the population to remain at an equilibrium of $K/2$. This happens when the population is fished with constant effort

$$F_{MSY} = \frac{r}{2}. \tag{1.14}$$

Open-access fisheries are those where the ability to harvest fish is unrestricted, with the right to fish free and available to all. In such a fishery, if the total revenue is greater than the total cost, then there is excess profit available in addition to the minimum profit required to fish, which is included in the costs of fishing as an opportunity cost. Because access to the fishery is free, in an open-access fishery either an existing fisher will increase their effort, or a new fisher will join the fishery, in order to take advantage of the available profits. Consequently in an open-access fishery the total fishing effort will be such that an equilibrium is reached where the total revenue obtained by all fishers is equal to the total cost (including opportunity costs). This is what Gordon (1954) refers to as a "bionomic" equilibrium. In

this type of fishery, $P = 0$ which means that the open-access level of fishing effort is

$$F_{OA} = \frac{B}{c}Y \tag{1.15}$$

Figure 1.2 depicts the relationship between total cost, total revenue and fishing effort in an unstructured model. The MEY, MSY, and open-access (bionomic equilibrium) solutions are marked on this schematic diagram. It is observed that the relationship among fishing efforts for these solutions is $F_{MEY} \leq F_{MSY}$, which corresponds to the observation from (1.8) that for fixed population biomass, increased efforts result in increased yields. Therefore fishing to maximise economic returns means fishing at a lower effort than the level which maximises the actual biomass attained (unless there are no costs of fishing, in which case MEY and MSY are equal). Because fishing depletes population levels (Baum et al., 2003), fishing with the higher effort F_{MSY} results in a lower equilibrium stock level than fishing with the lower effort F_{MEY}. Because MEY maximises the value of revenue minus costs, and the bionomic equilibrium occurs where revenues and costs are equal, the relationship $F_{MEY} \leq F_{OA}$ also holds, with these efforts only being equal in the trivial case where both are zero. There is no strict relationship between MSY and the bionomic equilibrium, with the relative position of F_{OA} and F_{MSY} depending on both the slope of the revenue curve and the cost per unit effort of fishing.

1.2 Game theory

One way to model the decision making behaviour of fishers is to treat the fishery as a game. Games are strategic situations in which one or more individuals interact with each other and their environment (Tadelis, 2013). Approaches of this type began with von Neumann and Morgenstern (Von Neumann and Morgenstern, 1944), who mainly considered zero-sum games, in which gains by one player must be equally matched by losses from another. Their work was famously built upon by Nash, who considered both non-cooperative (Nash, 1951) and cooperative (Nash, 1953) (non-zero-sum) games between two players, and who devel-

oped the important concept of a Nash equilibrium solution. In a game, each player faces a decision problem, or a choice between a set of possible actions. For each possible action that a player can take, they will experience a corresponding outcome. Players each have preferences for specific outcomes, and we can define a preference relation which ranks the order in which the player desires these outcomes (Tadelis, 2013).

It is possible to construct a payoff function, from the set of possible actions to the set of outcomes, which represents a player's preference relation. A payoff function allows us to directly consider a player's preferences for the particular action or actions which will maximise their own payoff, rather than having to refer to their preferences over outcomes (Tadelis, 2013). Each player is assumed to be a rational "Homo economicus" or economic human, who will always choose whichever action maximises their payoff function with respect to all possible outcomes. This requires the player to completely understand the decision problem, meaning they are aware of all possible actions, all possible outcomes, which outcomes will occur due to the selection of each action, and their own preferences for these outcomes.

In a game, each of the players independently faces a decision problem. Strategies define the behaviours, or decision making processes, a player can take in order to accomplish a specific goal. A strategy must specify exactly which action a player will take in any situation. Players play with pure strategies when they have a deterministic response to every situation. Stochastic responses are called *mixed strategies*. Markov strategies are those that depend only on the history of the game (in terms of state variables and players' actions), and can be either pure or mixed. A normal-form game requires three components: a finite set of players, a collection of sets of pure strategies, (that is, a set of possible pure strategies for each player), and a set of payoff functions, each of which give a payoff to the player for every possible combination of strategies (Tadelis, 2013). A strategy profile is a set containing a single strategy for each player, and so fully specifies the actions which will be taken in a game (Tadelis, 2013).

A significant finite normal form game is the prisoner's dilemma (Rapoport and Chammah, 1965). In the defining example, two players (the prisoners) have been arrested for a serious crime, but there is only evidence of a more minor crime being committed. If neither player is convicted of the serious crime, they will both be rewarded (Re) with a short sentence. The players are separated, and each is tempted (Te) to confess or defect, and implicate the other, which will reduce their sentence to an even shorter one. If one prisoner does this and the other does not, the prisoner who does not confess will receive the sucker's payoff (Su), which is an increased sentence. However, if both players confess, and so implicate the other, they will both be punished (Pu) with a longer sentence than if neither had confessed. It is important here that the sucker's payoff Su is a longer sentence than Pu, such that if a player will be motivated to defect in order to obtain Pu over Su. This game is symmetric, with the same actions and payoffs being available to each player. The possible actions are the choice to cooperate with each other (C), or to defect (D), and the payoffs are the length of prison sentence received. In the above example we consider the payoffs to be the negative of the length of the prison sentence, so that preferable payoffs are larger. For a game to have the structure of a prisoner's dilemma, the following inequality of payoff values must be satisfied:

$$Su < Pu < Re < Te \tag{1.16}$$

This game can be represented as the matrix shown in Table 1.1.

		Player 1	
		Cooperate (C)	Defect (D)
Player 2	Cooperate (C)	(Re, Re)	(Te, Su)
	Defect (D)	(Su, Te)	(Pu, Pu)

Table 1.1: The standard payoff matrix for a two-player prisoner's dilemma. Entries are payoff vectors, where the first entry is the payoff to player 1, and the second entry the payoff to player 2.

In the second chapter of this thesis, we show that the fishing behaviour of two players in a simple bioeconomic fisheries game is equivalent to a prisoner's dilemma, and evaluate the

relative effectiveness of different strategies on economic outcomes for each player, and for the abundance of the fish population.

An equilibrium of a game is one of the possible actions which will result in an outcome that is more likely than others (Tadelis, 2013). Equilibria must be self-enforcing, meaning that as players are in control of their own actions, they will only choose such an action if it is in their own best interest. In this context actions are considered in a player's best interest if they maximise the player's payoff, consistent with that player's beliefs about what is occurring in the game. The social optimality of outcomes can be compared by considering the Pareto optimality of the strategies that lead to them. A strategy is said to Pareto dominate another if it leads to an outcome which is better off for at least one player, without making any players worse off (Tadelis, 2013). If a strategy is not Pareto dominated by any other strategy profiles, then it is Pareto optimal. For example, in the above prisoner's dilemma example, the strategy profile (C, C) Pareto dominates the strategy profile (D, D) as both players will receive a strictly greater payoff when both cooperate, according to Equation (1.16). (C, C) is not Pareto dominated by strategy profiles where each player chooses a different action; if either player were to deviate from this profile by defecting (leading to the profile (C, D), it would cause the other player to receive the sucker payoff Su. By the strict ordering of payoffs in the prisoner's dilemma, that player receiving the sucker payoff will be worse off than if they had received the reward payoff Re. The strategy profile (C, C) is Pareto optimal, as there is no playable strategy in the prisoner's dilemma that Pareto dominates it.

Pareto optimality does not necessarily lead to the best outcomes for an individual player, or for the overall well-being of all players. A strategy s_1 is *strictly dominated* by another strategy s_2 if choosing s_2 results in a better outcome for the player regardless of which strategy the other player chooses. In terms of the prisoner's dilemma, D strictly dominates C, as it will result in a better payoff for Player 1 than choosing C, whether Player 2 plays

D or C. Rational players will never play a strategy that is strictly dominated by another strategy (Tadelis, 2013). Furthermore, if a game has an equilibrium that is a strictly dominant strategy profile, then this equilibrium is unique (Tadelis, 2013). A strategy s is a best response to a given strategy s' by the other player if it results in a better payoff than any other possible strategy. Strategies that are strictly dominated by another cannot be a best response (Tadelis, 2013).

An extremely important solution concept for non-cooperative games is that of the Nash equilibrium (Nash, 1951), a situation in which none of the players can benefit from changing only their own strategy. A Nash equilibrium occurs when every player is playing a best response to their own beliefs, and the beliefs of all players about their opponents' behaviour are correct (Tadelis, 2013). Any combination of strategies in which all players are playing mutual best responses to all other players' strategies is a Nash equilibrium (Tadelis, 2013). Consequently at a Nash equilibrium, no player can improve their payoff by unilaterally deviating from their own strategy. In the context of the prisoner's dilemma, (D, D) is a Nash equilibrium, where neither player could receive a better payoff by unilaterally changing their behaviour. To do so, the player's only allowable other choice of action is C; they would then receive the Su payoff, which is strictly less than the payoff Pu they receive at the Nash equilibrium (D, D). On the other hand, the Pareto dominant strategy profile (C, C) is not a Nash equilibrium, as the strategy D strictly dominates C, and so at (C, C) either player could benefit from an increase to their own payoff (as $Te > Re$) by individually changing their strategy to D. The Nash equilibrium of a non-cooperative game can be fully described by the set of strategies which leads to it, and their corresponding payoffs. In the prisoner's dilemma, the strategy profile (D, D) and the payoff vector (Pu, Pu) therefore constitute a Nash equilibrium.

Games can be characterised by the amount of information available to each player. In games of *complete information*, full knowledge about each of the other players is available.

This means that each player is aware of their opponents' utility functions, possible strategies, and payoffs (Tadelis, 2013). Consequently in games of incomplete information, a player is unaware of some information about the way other players will play the game. Games like the prisoner's dilemma, in which each player makes a decision at a single point in time, are called stage (or single-shot) games. When the same stage-game is repeated a number of times, it is called an iterated (or repeated) game (Tadelis, 2013). Iterated games can be either finitely repeated, or infinitely repeated. In an iterated game, players are able to choose an action at each iteration according to their strategy profile, in which specific choices can be made according to the entire history of the game, and not just its present state (Tadelis, 2013). Repeated games allow for further concepts of the information available to players. In particular, they allow the for the idea of *perfect information*. In games of perfect information, each player, at the point at which they are making a decision, knows all events which have previously occurred including all past actions by other players and any initial state of the game (Tadelis, 2013). It follows that in games of imperfect information, players do not know all past events. Games of imperfect information also represent situations in which players may have uncertainty about the effects of acts of Nature on the game (Tadelis, 2013). Iterated games also allow new solution concepts. A pure strategy profile is a *subgame perfect equilibrium* if its restriction to every subgame is a Nash equilibrium (Tadelis, 2013). A subgame is a subset of a game, that is a game in its own right, and in which the history of the game can be ignored. Iterated games provide a valuable tool for mathematically modelling multi-stage decision making.

1.3 Stochastic optimisation models of harvesting

Stochastic optimisation models are methods for maximising or minimising an objective function where there is randomness present in the optimisation problem. Markov decision processes are a class of stochastic optimisation models in which discrete time stochastic processes are used to model the multi-stage decision making behaviour of a single individual in

a stochastic environment (Puterman, 2014). In a Markov decision process, the individual makes a decision about which actions to take in each state, at each time, for a problem which may have either a finite or infinite time horizon. In this type of stochastic process, the system is in a particular state s_n at any time n. The action a_n chosen by the decision maker will cause the state to change (in a probabilistic manner) to a new state s'_{n+1} at the next point in time. Markov decision processes have the Markov property of memorylessness, meaning that the next state in time depends only on the current state and action, and not on any previous states or actions. A state transition matrix Q gives the probability $Q(a_n, s_n, s'_{n+1})$ of transitioning from s_n to s'_{n+1} when the action a_n is taken. The decision maker will obtain a specific outcome or reward $R(s, s', a)$ depending on the state transition and their own choice of action. The future value of rewards is discounted by a factor δ. Solutions Π prescribing the best choice of action for each scenario are called optimal policies, and are optimal in the sense of maximising the decision maker's total expected (discounted) reward over the entire time horizon if that policy is followed. That is, setting $a_n = \Pi(s_n)$ maximises

$$\sum_{n=0}^{\infty} \delta^n R(a_n, s_n, s'_{n+1}) \tag{1.17}$$

Optimal policies may not be unique, but there exists at least one for each Markov decision problem (Puterman, 2014). In order to determine the best choice of action in a certain state and at a certain time, the decision maker must pay regard to both the immediate reward obtained from choosing that action, and the potential value available in the future, depending on the resulting state. Optimal policies are called stationary when they converge in time, meaning that the best action in each state does not depend on time (Puterman, 2014).

Markov decision processes are commonly solved by using a backwards induction process called value iteration (Puterman, 2014). In this method, $V_n(s_n)$, the maximum discounted expected reward over the remaining time period at time n and in state s_n, is considered. The best decision at each period is found recursively by iterating backwards through time,

beginning with an initial value function V_{n+1}, which could correspond to the last time-point in a finite time horizon problem. In an infinite time horizon problem the initial value function could correspond to the value of some initial guess for an optimal policy, or could correspond to the situation in which only immediate rewards are valued for a fishing period. After initialisation, the next time period is then considered, with the action which maximises the present and future value for each state chosen as the new value function V_n. This process is then iterated until subsequent value functions V are sufficiently close together (Puterman, 2014). The recursive values are found by solving the Bellman equation (Bellman et al., 1954):

$$V_n(s_n) = \max_{a \in A} \{R(a, s_n, s'_{n+1}) + \delta \sum_{n=0}^{N} Q(a, s_n, s'_{n+1}) V_{n+1}(s'_{n+1})\} \tag{1.18}$$

An alternative way to solve Markov decision processes is policy iteration (Puterman, 2014). This method relates directly to the structure of Markov decision problems, and can be used for stationary infinite-horizon problems (Puterman, 2014). In general, policy iteration finds solutions less efficiently than value iteration. After an initialisation step, policy iteration solves Markov decision problems by iteratively performing a policy evaluation step, in which the value function for a specific policy is calculated, and a policy improvement step, in which a new policy for the next time period is chosen (Puterman, 2014). This process is repeated until policies converge through time. We will use Markov decision processes in Chapter 3 of this thesis to investigate the optimal way to harvest a fish population with noisy dynamics. This models the problem faced by a monopolist (i.e. a single agent) with sole access to a noisy renewable resource.

White (1985, 1993) provided a survey of applications of Markov decision processes, both in "real" cases in which results were actually implemented as part of a decision making framework (White, 1985), and in more hypothetical situations motivated by real world problems (White, 1993). These included references to a series of papers by Mendelssohn on the application of Markov decision problems to modelling the harvesting of salmon (Mendelssohn,

1978, 1980, 1982). These papers were an attempt to decide escapement levels of the fishery: how many members of this fish population should be left to reproduce at the end of a fishing period, and how many should be caught. Mendelssohn (1980) utilised a lumped parameter Ricker model of the fish population, based on the stock-recruitment relationship of sockeye Salmon in Alaska. Policies were found which maximise the total expected discounted yield in biomass; the economic cost of fishing was not considered, and as such payoffs were not in terms of the actual profit obtained from fishing. The optimal policy derived in this model takes a simple form: there is a threshold population level, below which there should be no fishing. For any population state above this threshold, the optimal action was to fish down to the threshold population state. Mendelssohn (1978) presented an analysis of a single-species, multi-age model of a fishery. Age dependent mortality, reproduction, and size were included in the population model. It was noted that it is difficult to define fishing behaviour which will meet traditional goals such as an MSY, a constant number of fish harvested, or a stable age distribution in an age-structured model. They presented two models, one of which was a generalisation of the Beverton-Holt population model, and the other the Ricker model. In the Beverton-Holt model, the cohorts are separable, and so optimal policies were found for each age group independently of the remainder of the population.

The exploitation of populations in which growth is dependent on population size was discussed by Mann (1970). Differential equations representing the growth of males and females in the population were developed, with a Markov decision process used to determine the yield which should be extracted of each sex, in each fishing period. Lane (1989) considered a Markov decision process model of fishing where the fish population dynamics are not directly observed, meaning the fisher did not have direct knowledge of the state of the population (Lane, 1989). The type of approach is known as a partially observable Markov decision process. Specifically, fishers can measure their own yield, which provides an indirect measurement of the actual abundance of the fish population. In this paper, Lane investigated the decision making behaviour of independent fishers within a fishing season in terms of their

fishing effort and their decision where to fish. Fish population dynamics were modelled as growing according to a Markov chain in which there are a discrete number of population states.

1.3.1 Stochastic games as models of harvesting

In stochastic environments where there are two or more decision makers affecting a system, Markov decision processes are not an appropriate model. Instead, Markov (equivalently stochastic) games can be used Littman (1994). Markov games are an extension of Markov decision processes which allow multiple decision makers to choose actions which will affect the state variable, and to each receive a payoff for their choice of action. Value iteration methods can also be used to solve Markov games (Littman, 2001). While all Markov games will have an optimal policy, they do not necessarily have a deterministic stationary optimal policy (Littman, 2001). Solutions of Markov games can occur as two different types of strategies, based on different types of information structure (Fudenberg and Tirole, 1991). With closed-loop strategies players' strategies at time t can be dependent on the history of play up to that point. That is, players can act on variables other than calendar time (Fudenberg and Tirole, 1991). On the other hand, open-loop strategies are strictly functions of time (Fudenberg and Tirole, 1991). This means, for example, that players playing open-loop strategies are not able to directly respond to deviations by other players from an equilibrium strategy, or to realisations of environmental noise due to acts of nature. (Fudenberg and Tirole, 1991). They can instead only respond to the current system state, which may be affected by the actions of all players. It is ordinarily simpler to find open-loop equilibria than closed-loop equilibria (Fudenberg and Tirole, 1991).

In the Chapter 5 of this thesis, we look at how Markov games can be used to model the exploitation by two agents of a fish population with stochastic dynamics, and particularly how the strength of the noise in the population dynamics affects the optimal fishing behaviour and long-term outcomes for the stock.

In Martin-Herran and Rincón-Zapatero (2005), the authors found a class of differential games of fisheries in which there are Markov perfect Nash equilibria that are also Pareto efficient. Markov perfect Equilibria are "profiles of Markov strategies that yield a Nash equilibrium in every proper subgame" (Fudenberg and Tirole, 1991). The class of solutions found in Martin-Herran and Rincón-Zapatero (2005) is of interest because Pareto efficiency means that a socially desirable (in terms of aggregate payoff) outcome has taken place, which for these particular competitive games will be a self-enforcing outcome because of the Markov perfect Nash property they also satisfy. The strategies considered were Markov, meaning they depend only on the current state of the game. Unlike the stochastic games we look at in Chapter 4 they consider specifically continuous time population dynamics represented by differential equations. They noted that Pareto efficiency of a Nash equilibrium is very rare. The authors found necessary and sufficient conditions for Pareto efficiency of a Markov perfect Nash equilibrium, using as the general case a two-player game where each player has a unique action variable, and there is a single state variable. They formulated the problem as a system of partial differential equations, for which the solutions were the Markov perfect Nash equilibria. Chiarella et al. (1984) found examples of fishery games in which there are Pareto efficient Nash equilibria. Here players were only able to use open-loop strategies, meaning that these Nash equilibria are not subgame-perfect. In this fishery model each player in the game had indirect commonality of access to the shared fishery resource. This means that each player can harvest the shared stock when it is in their own exclusive economic zone, but not when fish stock moves across international boundaries.

Sobel (1981) presented sufficient conditions for "myopic" solutions of both (single-player) MDPs and mutli-player stochastic games. These solutions are those that can be obtained from optima of a static game. The condition that transition probabilities are dependent only on action and not present state do not apply to the games that arise from our models, and so the existence of myopic solutions is therefore not guaranteed. There was reference

to some forms of games in which myopic equilibria occur under some conditions, including a stochastic fisheries model in which there are multiple interacting species. This particular model was discussed in (Sobel, 1982). They presented a multi-species, multi age-class model of fishery, motivated by both biological reasoning (including temporal fluctuations in age structure, inter-specific predation and competition for prey) and operational reasons such as the landing of bycatch. Sobel (1982) acknowledged the difficulties, which also apply to our model, of using estimates of population biomass at the end (or start) of a fishing season as the decision variable. Such estimates are not necessarily accurate due to a lack of sampling data. This model also assumed that natural mortality is not large during the fishing season.

McKelvey (1997) contextualised multinational (transboundary) fisheries management as involving brief and intense fish wars characterised by withdrawals from cooperative agreements, economic downturns in fishing communities, and heavily depleted fish stocks. Non-cooperative game theory offers a method to understand such conflicts in commercial fisheries. They noted that uncertain and asymmetric information can have implications for harvesting models of fisheries. In our fishery models we assume symmetric information for players. McKelvey (1997) developed a simple harvest model of a seasonal "interception" fishery in which a migratory transboundary stock was harvested by two distinct agents with their own regulatory zones. This was a sequential model in which the first fishing fleet initially harvests the fish population. The stock then migrates to the second zone, where the second fleet has unrestricted access to solely exploit the stock. Finally, the fish stock migrates back to the initial zone, where the first fleet again is solely able to harvest the stock. Like our approach, first a deterministic, then a stochastic version was analysed. The stochasticity in this model occurred entirely in the payoff functions of each player, with deterministic population dynamics remaining. They used memoryless strategies, with behaviour only taking place as a reaction to the present level of the fish stock (in contrast with Kaitala and Pohjola (1988) who used a cooperative differential game model in which all past deviations from agreement were memorised). The population model used by McKelvey (1997) described the harvest

and escapement of a stock in which there was no natural mortality, with escaping stock spawning recruits for the next season's population. They assumed the unit price of fish to be constant regardless of harvest size, and that the unit cost of harvest rises as population declines, equivalent to a fixed cost per unit effort. The problem was solved using optimal control methods. It was found that open access to the fishery drives stock levels down to the break-even level of payoff, which includes the opportunity cost of fishing. The optimal escapement stock level was found to be the point where the marginal gain in revenue that would be obtained from a decrease in escapement (that is, an increase in harvest) was equal to the loss in the marginal value of recruitment to the next season's harvest.

Chapter 2

Emergence of balanced harvesting from size-spectrum models of a small-scale fishery

2.1 Introduction

Fish are an essential renewable resource throughout the world, both for the purposes of food and employment, and so fish stocks should be protected (FAO, 2014). In commercial fisheries, large fish typically have a greater economic value than small fish, given that they reach higher market prices per unit biomass (Sethi et al., 2010; Tsikliras and Polymeros, 2014). However, in other areas of the world where there is poverty and malnutrition, the economic value of fish is less important than its actual food content (in terms of the total biomass caught) (Beveridge et al., 2013; FAO, 2014). In poor areas with small-scale fisheries, small fish are as valuable as larger fish on a per-biomass basis, as they can be more easily cooked (Kawarazuka and Béné, 2011; Longley et al., 2014).

Existing high-level fisheries' management strategies have aims which include protecting small fish (Botsford et al., 1997), with the goal of letting every individual fish have the opportunity

to spawn (Froese et al., 2015). However, methods which involve features such as minimum target sizes can have unintended negative consequences on fish stocks, including disruption of size distributions and fisheries' induced evolution, as well as having ecosystem level impacts on non-target species. Balanced harvesting, in which total fishing mortality is broadly distributed across species and individual sizes in proportion to natural productivity, has been proposed as a potential alternative approach, which could mitigate the negative side-effects of fishing on marine ecosystems (Garcia et al., 2012).

The choices made by individual fishers about how to fish can be modelled as a non-cooperative game (Tadelis, 2013). Each fisher, or player, faces a decision problem in which they must choose between a series of actions, each of which have a corresponding outcome. Players in a game act to optimise their choice of action in order to obtain the best possible outcome. Static games occur when each player simultaneously and independently chooses their respective actions (Tadelis, 2013). This condition means that all players make their own specific decisions without knowing which actions any other players are taking. In a static game, outcomes are then determined for each player based on the combination of all players' actions. Any combination of strategies in which all players are playing mutual best responses to all other players' strategies is a Nash equilibrium (Tadelis, 2013). Consequently at a Nash equilibrium, no player can improve their payoff by changing their own strategy unilaterally.

Fish populations can be modelled by considering their size-spectrum, the abundance of all members of the population of a certain size over the size range from eggs to adults, (Law et al., 2012) which occupies several orders of magnitude. Dynamic size-spectrum models of a population explicitly allow the abundance of the population over the whole size-spectrum to be tracked over time. A schematic diagram representing a size spectrum model of a single fish species supported by a plankton producer spectrum is depicted in Figure 2.1. Such models track biomass levels to determine the growth, birth, and death of organisms of a certain size in the population (Law et al., 2012). One way to model the size-spectrum of a population is

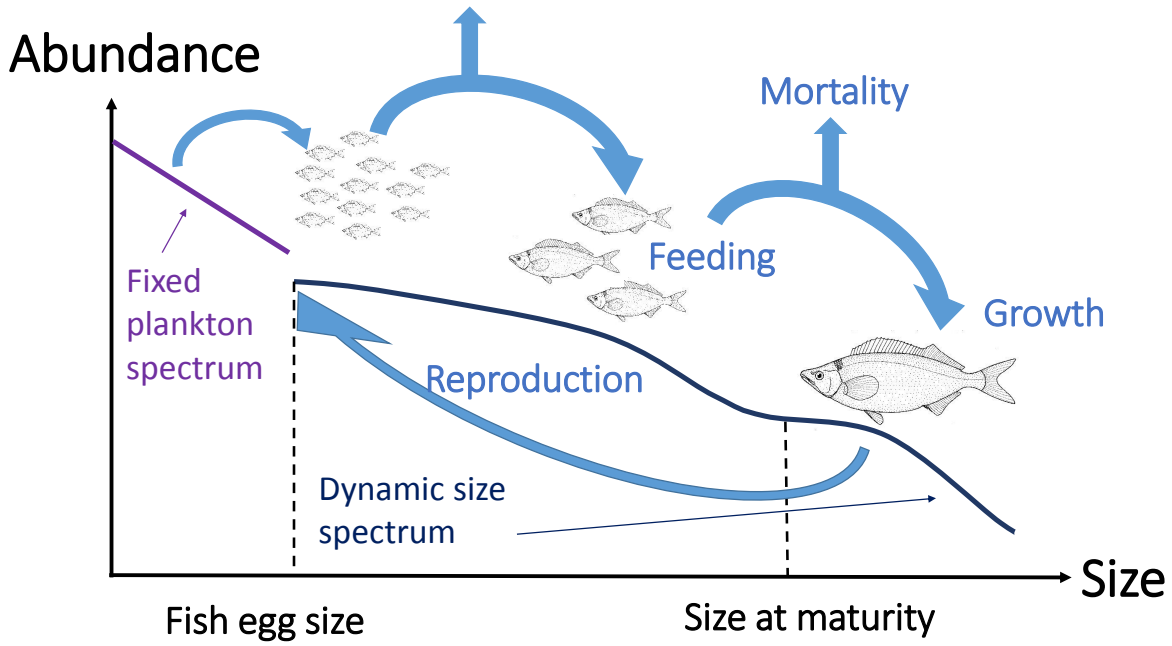


Figure 2.1: A schematic diagram of a single species size spectrum model, highlighting the different processes by which biomass is transferred through the ecosystem.

to use a McKendrick-von Foerster equation. This is a size-structured, continuous time PDE, which has the biomass density of the population at size x and time t as its solution (Kot, 2001).

In this chapter, we will combine a size-spectrum model of a single fish species with an individual-based model of the behaviour of a large number of fishers, each trying to exploit a small-scale communal access fishery. In this type of small-scale fishery, fishers may have to be able to choose from different fishing methods and equipment, resulting in different size selectivity. In this model, we assume they will do this by changing their net mesh size. We investigate how the size-selectivity behaviour of fishers changes based on relative profits available from the fishery, allowing for fishers to enter or exit the fishery depending on whether potential gains meet their expectations, accounting for the opportunity cost of fishing. We relate the emergent behaviour of fishers (with regard to their aggregate size-selectivity) to the concept of a Nash equilibrium, and show that the fishing behaviour represents a move

towards balanced harvesting of the fish population.

2.2 Size spectrum model

We consider a dynamic size spectrum model of a population of a single fish species living in a non-seasonal environment as used by Law et al. (2012). This model allows for bookkeeping of biomass as it is transferred throughout members of the population due to biological processes of growth, reproduction, and mortality in order to obtain population levels. These biological processes are calculated dynamically within the model, and are thus not externally specified. Life histories of fish are described according to their size, w, with the abundance of fish at that size given by $\phi(w)$. As it is biologically infeasible for fish to continue growing forever, they are considered to have an asymptotic maximum weight w_∞. Similarly, fish are assumed to have a minimum size w_0 corresponding to the size of a fish egg, which is assumed to be of constant size.

The model is expanded by including a producer spectrum in the form of a plankton community, which is assumed for simplicity to be fixed. The fixed plankton spectrum provides a continuously present source of prey for the young members of the fish population. The relationship between plankton abundance $\phi_p(w)$ and body size w is considered to be that of a power-law $\phi_p(w) = u_0 w^{-\gamma}$ with exponent $-\gamma = -2$, and with an upper bound on plankton size set at $w_{0,max} = 0.02$g, in order to match empirical observations of similar environments (San Martin et al., 2006).

2.2.1 Fish growth

The growth characteristics of individual members of the fish population are investigated by considering the exchange of biomass within the fish population, and between members of that species and the producer spectrum. Fish grow by preying on smaller organisms in order to accumulate biomass. In this single-species model, predation is predominantly conspecific,

and so the trophic chain is comprised of that species and the producer spectrum. Large fish prey on smaller fish, and small fish prey on plankton. The feeding preference of fish of size w for prey of size w' is given (as a function of their size ratio) by the feeding kernel

$$s\left(\frac{w}{w'}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left(\log(w/w') - \beta\right)^2}{2\sigma^2}\right). \tag{2.1}$$

In the feeding kernel s, β represents the mean preferred log ratio of predator to prey size, and σ the standard deviation of this ratio. These parameter values are motivated from stomach contents data (Ursin, 1973). It is assumed that fish never prey on any larger individuals, such that when $w' > w$, $s(w/w') = 0$. The spatial region occupied by the population is assumed to be well-mixed, so that the position of any fish has no bearing on its likelihood of interacting with any other fish. Fish move through the spatial region with a search rate dependent on their size, and parameterised by an exponent α and a feeding rate constant A (with units $\mathrm{m^3 y^{-1} g^{-\alpha}}$). This search rate Aw^α, gives the volume V searched per unit time, and represents the ability of a fish of size w to prey both conspecifically, and on the producer population. There is an inefficient conversion of consumed prey biomass into predator biomass: an ecological conversion efficiency κ is used to account for losses due to feeding, digestion, and the metabolic cost of growth. That is, a proportion κ of consumed prey biomass is assimilated into predator biomass. Consequently the mass-specific rate of biomass assimilation by a predator of size w is given by

$$g(w,t) = A\kappa w^{\alpha-1} \int_{w0}^{w_\infty} w' s\left(\frac{w}{w'}\right) \left(\phi(w',t) + \phi_p(w')\right) dw'. \tag{2.2}$$

2.2.2 Reproduction

At the beginning of their lives, small fish allocate all of the usable biomass accumulated from feeding to somatic growth. Fish mature and gain the capacity to reproduce as they grow to a certain size. This maturing process is modelled by an ogive with a midpoint at size w_m, and a transition width ρ_m. As fish mature, they begin to allocate an amount of assimilated

biomass away from growth and towards reproduction. We define the function representing the allocation of energy to reproduction as

$$E(w) = 1 - \left(\frac{w}{w_m}\right)^{\rho} \left[1 + \left(\frac{w_m}{w}\right)^{\rho_m}\right]^{-1}. \tag{2.3}$$

The reproduction function is parametrised by an exponent ρ, chosen to be equal to $\gamma - \alpha - 1$, so that the fecundity of a mature fish will be approximately constant with respect to weight. $E(w)$ gives the proportion of usable biomass allocated to somatic growth, with the remaining $1 - E(w)$ allocated to reproduction by a fish of size w. The proportion of biomass allocated to growth decreases with fish size after maturity has been reached. Fish who have reached w_∞, the asymptotic maximum weight, allocate all of their usable biomass gained by feeding to reproduction, and none to somatic growth, and thus represent idealised BOFFFFs in the size-spectrum model. The size of fish eggs is distributed according to a birth kernel $b(w)$, chosen to be a Dirac$-\delta$ function such that all fish eggs are of the same size, w_0. The total rate at which reproductive biomass is created in the population is given by the function

$$R_b(t) = \int_{w0}^{w_\infty} (1 - E(w))\phi(w)g(w,t)w dw. \tag{2.4}$$

The biomass output due to reproduction is assumed to be immediately converted into individuals of the fixed egg size w_0.

2.2.3 Fish mortality

We include several sources of mortality for fish. As detailed above, fish can die by being preyed on by larger fish of any size. We let $\mu_p(w,t)$ be the mortality rate of fish of size w due to predation. The rate at which mortality due to predation occurs is calculated by integrating the abundance of predators of size w', the feeding kernel, and the predator search rate:

$$\mu_p(w,t) = A \int_{w0}^{w_\infty} w'^{\alpha} s \left(\frac{w'}{w}\right) \phi(w',t) dw' \tag{2.5}$$

Intrinsic mortality $\mu_i(w,t)$ gives the probability for a fish of size w to die of natural causes other than predation, and for small fish is assumed to be small compared to predation mortality, and increasing with body size. Intrinsic mortality, which is parametrised by the exponent ξ, is calculated as

$$\mu_i(w,t) = \mu_0 w^{-\xi} \frac{g(w_0,t)}{g(w,t)}. \tag{2.6}$$

If the stock is being exploited, organisms will also die due to being harvested, and so we include fishing mortality $\mu_f(w,t)$. For ease of reading, we omit the time arguments in all mortality and growth functions, writing for example $\mu_f(w)$ and $g(w)$. The total mortality rate due to all causes is

$$\mu(w) = \mu_p(w) + \mu_i(w) + \mu_f(w). \tag{2.7}$$

2.2.4 Population level dynamics

Individual level growth, reproduction, and mortality processes are scaled up to ecosystem level dynamics through the use of the McKendrick-von Foerster equation

$$\frac{\partial}{\partial t}\phi(w) = -\mu(w)\phi(w) - \frac{\partial}{\partial w}\left(E(w)g(w)\phi(w)w\right) + \frac{b(w)R_b}{w}. \tag{2.8}$$

Here for simplicity we have withheld arguments of time t. The parameter values used in this size-spectrum model are given in Table 2.1, and have been chosen to represent the life histories of African catfish (*Clarias gariepinus*). Steady state abundances for an unexploited population were obtained by running the size spectrum model with not fishing mortality (that is with $\mu_f(w) = 0$). The biological processes $g(w)$ and $\mu(w)$ are determined by numerically integrating over the range of possible prey or predator body sizes. The steady state population was used as an initial condition in simulations where the dynamic size-spectrum was coupled with an individual-based fishing model, which is described in the next section of this chapter.

Parameter	Comments	Value	Units
w_0	Fish egg mass	0.001	g
w_m	Mid-point of maturation ogive	2200	g
w_∞	Asymptotic maximum weight of fish	2290	g
A	Feeding rate coefficient	750	$\mathrm{m^3 y^{-1} g^{-\alpha}}$
α	Search rate scaling exponent	0.8	-
ξ	Intrinsic (non-predation) mortality exponent	0.15	-
κ	Conversion efficiency of biomass intake	0.2	-
β	Feeding kernel mean	5	-
σ	Feeding kernel standard deviation	2.5	-
k_F	Individual harvesting intensity	0.0025	$\mathrm{y^{-1}}$
β_i	Mean target size of ith fisher's selectivity strategy	Variable	-
θ	Standard deviation of size-selectivity function	0.2	-
ρ	Reproduction function exponent		-
ρ_m	Width of transition to maturity		-

Table 2.1: Parameters chosen to represent the life histories of African catfish (*Clarias gariepinus*) used in the size-spectrum model to calculate population dynamics. The length at first maturity is approximately 30.8 cm (Kolding et al., 2003) (Fishbase, www.fishbase.org/summary/1934), with asymptotic length 67.5cm. Length l is converted to mass w using $w = al^b$ with $a = 0.008\mathrm{g\ cm^{-b}}$ and $b = 2.983$ (Kolding et al., 2003). Other parameters values are as in Law et al. (2016).

2.3 Individual based model

We extend an existing (Law et al., 2012) individual based model of the exploitation of a single species, which considered a fixed number of fishers exploiting a population, to allow for a dynamic number of fishers. The dynamic size spectrum model of a marine population was coupled with an individual-based model of N fishers in a small-scale fishery. Each fisher acts with behaviour determined according to a fixed, identical set of constraints. Individual fishers act with the goal of maximising their own personal profit in the short term. Consequently they do not consciously collude in order to harvest in a sustainable manner over the long term, and completely discount future profits. Here for ease of calculation, and to show clearly the relationship to individual size-selectivity, we refer to the yields Y_i obtained by fishers rather than their profits P_i. Note however that there is a one-to-one relationship $P_i = aY - c$ between yield and profit for a fisher with a fixed fishing effort.

2.3.1 Fishing methods

Fishers are considered to have a limited harvesting ability, representing restrictions including the actual capacity of the nets used, or the number of available boats. Each fisher is able to harvest fish according to an individual size-selectivity function, which takes the form of a Gaussian distribution over log body mass. As a unimodal distribution, this Gaussian selectivity function represents a harvesting method such as gillnet or hook fishing being used by the fisher. The amplitude of each selectivity function is given by k_F, which represents the intensity of a single fisher's harvesting effort, corresponding to the actual capacity of the fishing gear. Each of these individual fishing functions is also characterised by the parameters θ, the standard deviation of the individual size selectivity, which is constant for all fishers, and β_i the mean size targeted by the ith fisher, corresponding to the choice of net mesh sizing. Thus the ith fisher has an individual fishing function defined by

$$F_i(w) = k_F \exp\left(-\frac{(\log\left(\frac{w}{w_0}\right) - \beta_i)^2}{2\theta^2}\right). \tag{2.9}$$

The individual size-selectivity functions for each fisher therefore have identical area on the log body mass scale. Initially, each of the β_is is obtained from a uniform random distribution over the entire possible range of log body mass. The total harvesting pressure on the fish population is obtained as the sum of all of the individual fishing functions as

$$\mu_f(w) = \sum_{i=1}^{N} F_i(w). \tag{2.10}$$

This total size-selectivity function is a size-dependent population death rate, and can therefore be incorporated into the dynamic size spectrum model as $\mu_f(w)$, the fishing component of the death rate. The biomass yield obtained by the ith fisher is calculated by multiplying their own size-selectivity function with the population's biomass spectrum, integrated over all body sizes

$$Y_i = \int_{w0}^{w_\infty} wF_i(w)\phi(w)\,dw. \tag{2.11}$$

Therefore, the total yield is

$$Y = \sum_{i=1}^{N} Y_i. \tag{2.12}$$

2.3.2 Harvesting behaviour

Fishers periodically have the opportunity to change their individual size-selectivity behaviour, representing for instance a change in the net-mesh size used to catch fish. After every time period of length Δt, each fisher can change their size-selectivity strategy. The probability of such a change occurring for the ith fisher depends on the size of their own yield Y_i relative to the maximum yield obtained by any fisher. The probability that the ith fisher will change their value of β_i at the end of one such time period is given by

$$P = 1 - \frac{Y_i}{\max_{1 \leq j \leq N} Y_j}. \tag{2.13}$$

Individuals are considered to have a limited knowledge of their competitors. Specifically, over a given time period, they are aware of the maximum yield obtained by any fisher in the ecosystem, but do not have knowledge of the specific targeting strategy used to obtain that yield. Furthermore, they do not know the current state of the population size-spectrum, and they have no memory. Therefore while fishers have the ability to stochastically change their harvesting behaviour, they cannot directly do so to actively maximise their current yield. A change in harvesting behaviour occurs as a new mean target size β_i being selected for the relevant fisher. As fishers do not have knowledge of the optimal target size, β_i is again selected randomly from a uniform distribution over log body mass.

2.3.3 Economic motivations

The existing (Plank et al., 2016) individual based model of fishing was extended to allow a dynamic number of fishers in the fishery. As there is an economic cost associated with the act of fishing, each fisher will now require a certain level of revenue in order for them to remain active in the fishery. One common approach to directly modelling the population

dynamics under exploitation is to use the Gordon-Schaefer model (Schaefer, 1954; Gordon, 1954), which considers both the costs of fishing and the revenue obtained. Under the Gordon-Schaefer model, the rate of change of total fishing effort F is proportional to the net profit from fishing

$$\frac{dF}{dt} = k(aY(t) - c).$$

where a is the market price of fish per unit mass, $Y(t)$ is the aggregate yield at time t, c is the cost of fishing per unit effort and k is a constant of proportionality defining how rapidly fishing effort changes in response to profits. If net profit is positive, the total fishing effort will increase; if profit is negative, fishing effort will decrease. At equilibrium, the yield $Y(t)$ is given by the ratio c/a of unit cost to unit price.

We applied this model at the level of individual fishers to determine the rate at which fishers enter or leave the fishery. Recall that the fishing effort for each fisher is assumed to be constant, k_F, so the net profit for fisher i is $aY_i - ck_F$. We assume that if the average net profit is positive, more fishers will enter the fishery, resulting in an increase in total fishing effort. If average net profit is negative, fishers will leave the fishery, resulting in a decrease in total fishing effort. Note that this defines a critical yield $Y^* = ck_f/a$: if the average fisher's yield $\bar{Y}$ is above this critical yield Y^* then new fishers will enter the fishery and vice versa if $\bar{Y} < Y^*$. We use a stochastic model for the number of fishers entering or leaving the fishery at each time step. Specifically, the number of fishers entering or leaving during a time step of length δt is assumed to be a Poisson random variable with mean

$$\lambda = kN \left| \bar{Y} - Y^* \right| \Delta t$$

Consistent with the Gordon-Schaefer model, the rate of change of total fishing effort is proportional to the average net profit. For the numerical calculations in this chapter, we used values $k = 2.5 \times 10^{-7}$ and $\Delta t = 0.0137$y (that is, each fishing period is 5 days).

2.4 Results

2.4.1 Fixed number of fishers

We start by running the model with a fixed number of fishers, as in (Plank et al., 2016). The initial condition for the population size-spectrum was set to be the steady-state abundance when no exploitation ($F = 0$) occurs. The initial condition for the total fishing effort F was randomised by selecting each individual target log body mass β_i from a uniform distribution $U(\log(w_0), \log(w_\infty))$ over possible log body masses. The McKendrick-von Foerster equation ((2.8)) was solved using the method of lines (Schiesser, 2012), such that it was discretized in the spatial variable w, leaving derivatives in time t continuous. This resulted in a system of ordinary differential equations which were solved using the "ode15s" solver in MATLAB. After each fishing period of length δt, fishing behaviours were allowed to change, with new target sizes β_i allowed to be randomly selected for eligible fishers according to the process described in Subsection 2.3.2. The resultant total fishing effort F was calculated, and with (2.8) then solved for the subsequent fishing period, with this process repeated for a number of iterations.

Figure 2.2 (a) shows that after time $t = 100$ years, fishers respond to their objective of maximising their own yield by shifting their harvesting behaviours in order to target sizes in which fish have high biomass. Consequently, without explicitly constraining fishers with respect to their aggregate fishing behaviour, we observe that under the conditions of this model they will self-organise such that it appears that partially balanced harvesting over body size will occur. There is a peak in the aggregate fishing mortality near the 100g body size, the same size at which the maximum value occurs in the unexploited biomass spectrum (the blue solid line in Figure 2.2 (b)) This unexploited biomass is the population size-spectrum in the absence of fishing; in this case the fishing mortality rate $\mu_f(w) = 0$. In contrast, the exploited. Furthermore, in Figure 2.2 (b) we observe that the exploited biomass spectrum (red dashed line) is flattened over the range of targeted body sizes, such that the population

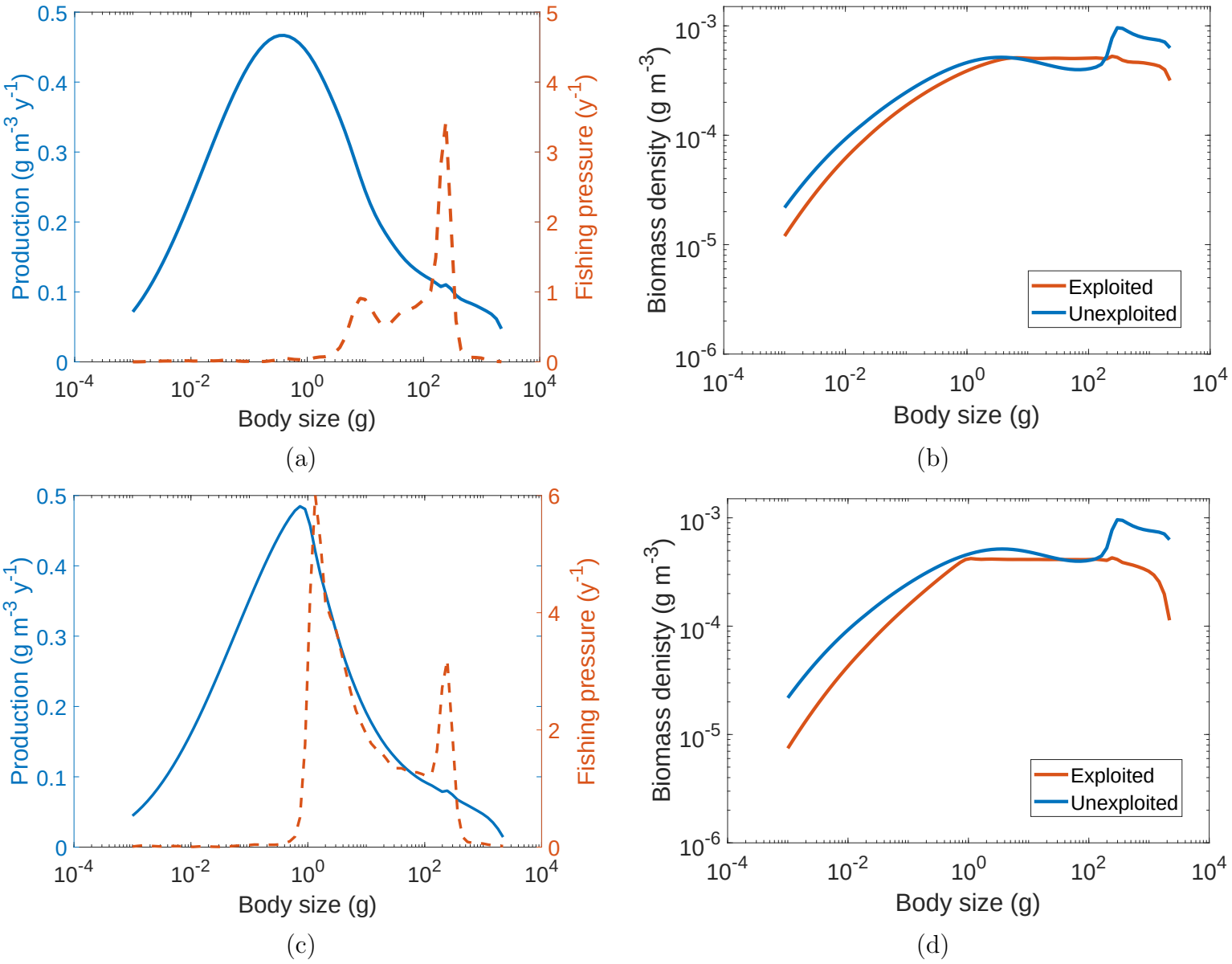


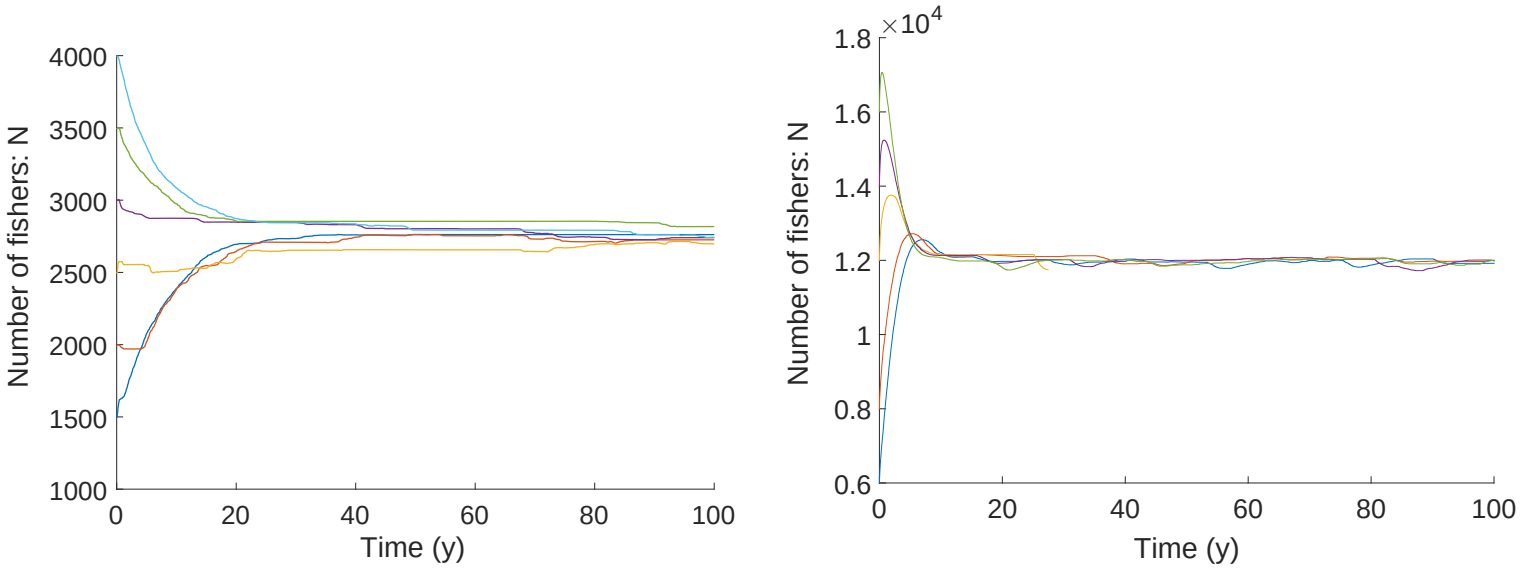
Figure 2.2: (a) shows that for a fixed number of fishers, regions of high exploitation (blue solid line) occur in regions of high productivity (red dashed line). In particular, peaks in the total fishing rate for a fixed number of fishers ($N = 2000$) occur at similar sizes to peaks in the population productivity spectrum. From (b) it can be observed that fishing causes a flattening of the biomass spectrum in the exploited size range (blue line) when compared to the unexploited biomass (red line).With a fixed number of fishers ($N = 2000$), the targeted region of the population has a constant biomass spectrum, meaning that fishers cannot increase their yield by changing their harvesting strategy. (c) shows that with a greater number of fishers ($N = 6000$), the total fishing rate (blue solid line) becomes closer to being proportional to productivity (red dashed line). With a fixed $N = 6000$, the relationship between total harvesting pressure and the productivity of the species becomes more apparent, showing that an emergent balanced harvesting situation has occurred. Furthermore, (d) shows that as the (fixed) number of fishers is increased, the width of the flattened part of the biomass spectrum under fishing pressure (red line) becomes larger. For $N = 6000$, a greater range of fish sizes are targeted, representing a greater breadth of exploitation of the ecosystem.

biomass is effectively constant over log body mass. The exploited biomass spectrum here is the fish size-spectrum after $t = 100$ years, under the pattern of fishing mortality $\mu_f(w)$ determined by our harvesting model. Consequently, if individual fishers attempt to change their fishing strategy by targeting another size range, they will not be able to increase their yield. This is because individual yield is proportional to both population biomass, which is here constant with log body size, and fishing intensity, and each fisher has a size selectivity function with identical area. Therefore, a Nash equilibrium has emerged in the system, in which no fisher can unilaterally alter their behaviour (by changing their target size) without obtaining an equal or worse yield for themselves. In Figure 2.2 (b), with $N = 2000$, the flattened region of the biomass spectrum is found for body sizes ranging from $\sim$ 8g to $\sim$ 200g.

In Figure 2.2 (c) and (d) we observe the effects of an increased number of fishers in the fishery. Figure 2.2 (d) depicts a flattened part of the biomass spectrum with greater width, meaning that as the number of fishers in the fishery is increased the range of exploited body sizes increases. Specifically, with $N = 6000$ fishers, the flattening of the biomass spectrum is observed for body sizes ranging from $\sim$ 1.5g to $\sim$ 200g. Furthermore, as the biomass spectrum is still constant over the harvested range, we note that the Nash equilibrium still exists for fishers when there is increased exploitation. Figure 2.2 (c) shows an increased correspondence between the natural productivity of the population and the size-dependent fishing pressure. The aggregate fishing function is now approximately proportional to productivity over the rest of the range of exploited body sizes. Consequently, we observe that a situation similar to fully balanced harvesting is occurring with this increased fishing pressure, as was expected from previous results (Plank et al., 2016).

2.4.2 Dynamic number of fishers

The numerical methods in the dynamic model were similar to that in Subsection 2.4.1, except that after every fishing period a random number of fishers were able to enter or exit the fishery, as described in Subsection 2.3.3. Figure 2.3 shows that regardless of the initial number



(a) (b)

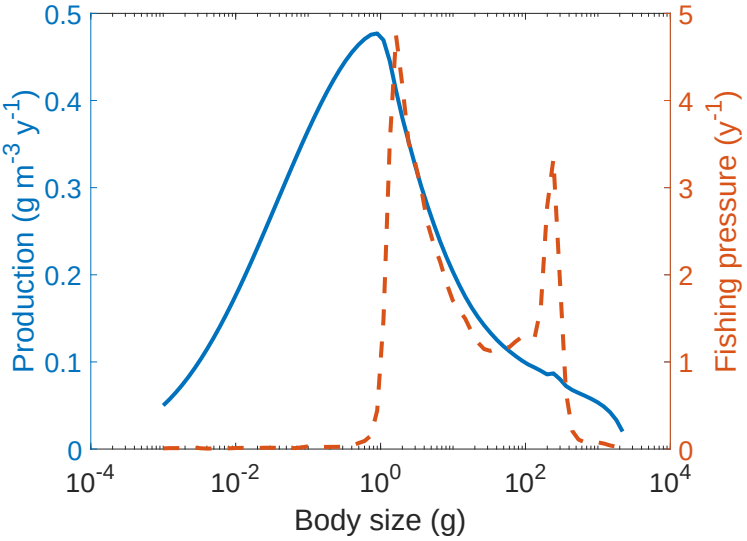
Figure 2.3: For any initial choice of fishers N_0, the number of fishers will converge to a stable equilibrium solution. The actual long-time number of fishers N depends on the size of the critical yield value Y^* required to keep individual fishers active in the fishery. (a) has a relatively high choice of $Y^* = 6\times10^{-6} g\, m^{-3} y^{-1}$, the number of fishers will converge over time to approximately 2800. (b) shows that for a smaller critical yield ($Y^* = 3 \times 10^{-6} g\, m^{-3} y^{-1}$), the total number of fishers after a long period of time will be greater, with N converging to around 12000 over time.

of fishers, the number of active fishers will converge over time to a level based on the critical yield for those fishers. While initial conditions do not affect the final number of fishers, and therefore do not affect the total level of exploitation of the fish species, the critical yield Y^* does. The number of active fishers which can be supported by the population in the long term decreases as the critical yield required to sustain these fishers increases. Figure 2.4 shows that over time fishers will self-organise into a Nash equilibrium solution, as in the fixed N model. For both high and low critical yields, the biomass spectrum is constant over the exploited range of body sizes. In Figure 2.4 (b), we see that the overall biomass density is similar to the unexploited state in the case of a high critical yield ($Y^* = 5.37 \times 10^{-6} g m^{-3} y^{-1}$), with a flattened region ranging over body sizes $\sim$ 2g to $\sim$ 120g. In Figure 2.4 (d), with greater fishing pressure (and a lower critical yield: $Y^* = 3.5 \times 10^{-6} g m^{-3} y^{-1}$), biomass density is lower than the unexploited state at each body size. However, the Nash equilibrium is still

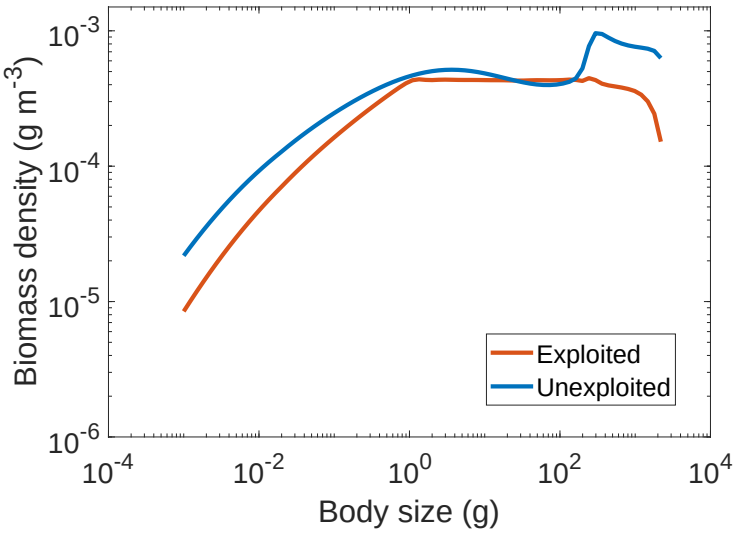
occurring with a greater number of fishers, and so no single fisher can alter their harvesting strategy in order to increase their personal yield. With this critical yield, the biomass spectrum is flattened over body sizes ranging from $\sim$ 0.5g to $\sim$ 200g. Recall from Section 1.2 Nash equilibria in a game can be fully described by the combination of all strategies taken by all players, and the payoffs obtained by each player. In the context of our dynamic fishing model, we see that the payoff that each player will individually receive is the critical yield. The strategies used by players are their individual fishing behaviours in terms of target size selectivity, with the overall fishing pressure representing a proxy for the full strategy profile.

Figure 2.4 (c) shows that as in the previous individual based model, the total fishing pressure is approximately proportional to the species' productivity over the total range of exploited sizes when harvesting is at a high level. When the total level of exploitation is lower, as in Figure 2.4 (a), the maximum in the total size-selectivity occurs at the same body size at which there is a peak in the unexploited biomass spectrum, again as in the model with a fixed number of fishers.

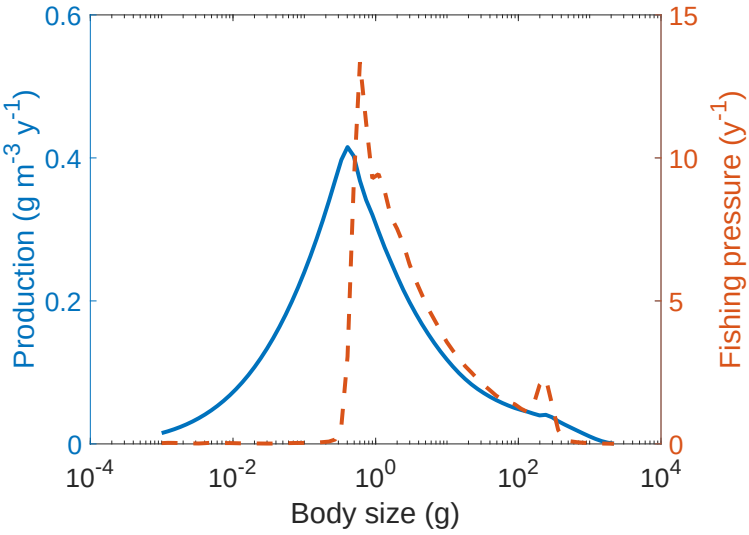
For a given critical yield, there is a specific number of active fishers that the fishery is able to support, who each obtain that yield. By considering a range of Y^* values and determining the number of fishers which can sustainably remain in the fishery, the total rate at which biomass is removed from the population (that is the total sustainable yield) can be determined. Figure 2.5 shows that a peak in the total sustainable yield occurs when the critical yield is set to a value of approximately $3.5 \times 10^{-6} \mathrm{g\,m^{-3}y^{-1}}$. This suggests that for critical yield levels above this value, the resource is being underutilised, as a greater amount of biomass could be removed from the population each year without irreversibly depleting the stock. Overfishing will occur for critical yields less than this value, with a larger number of active fishers causing the fish population to be depleted.



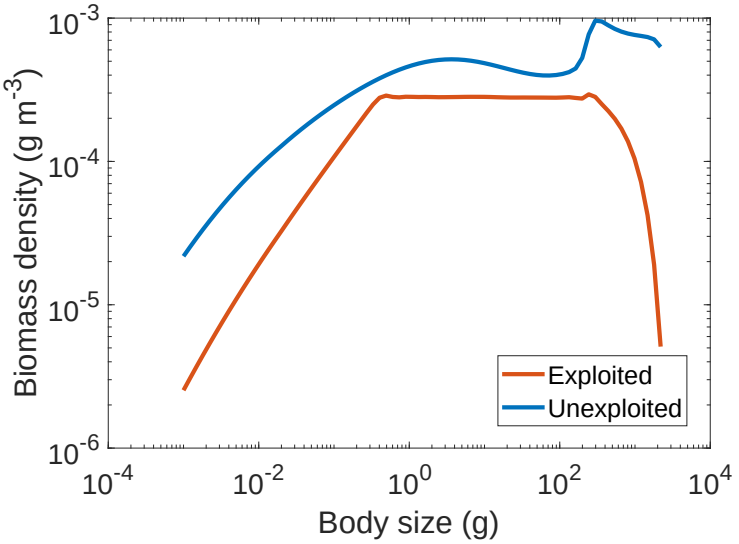
(a)



(b)



(c)



(d)

Figure 2.4: For a high critical yield ($Y^* = 5.37 \times 10^{-6} gm^{-3}y^{-1}$), there is a small flattened region of the biomass spectrum under exploitation (red dashed line) when compared to the unexploited biomass spectrum (blue solid line). (a) shows that high fishing pressure (red dashed line) occurs at sizes where the population has high productivity (blue solid line). A local peak in the overall size-selectivity function occurs at the body size at which the unexploited biomass spectrum has a maximum value. For a lower critical yield ($Y^* = 3.5 \times 10^{-6} gm^{-3}y^{-1}$) and thus higher exploitation level, the flattened section of the exploited biomass spectrum (red dashed line) in (d) is wide. Furthermore, actual biomass density levels are lower than the unexploited state (blue solid line) for all body sizes. At high levels of exploitation, as in (c), the total size-selectivity (red dashed line) is approximately proportional to productivity (blue solid line) over the entire range of relevant body sizes. Consequently at this level of harvesting, balanced harvesting is occurring.

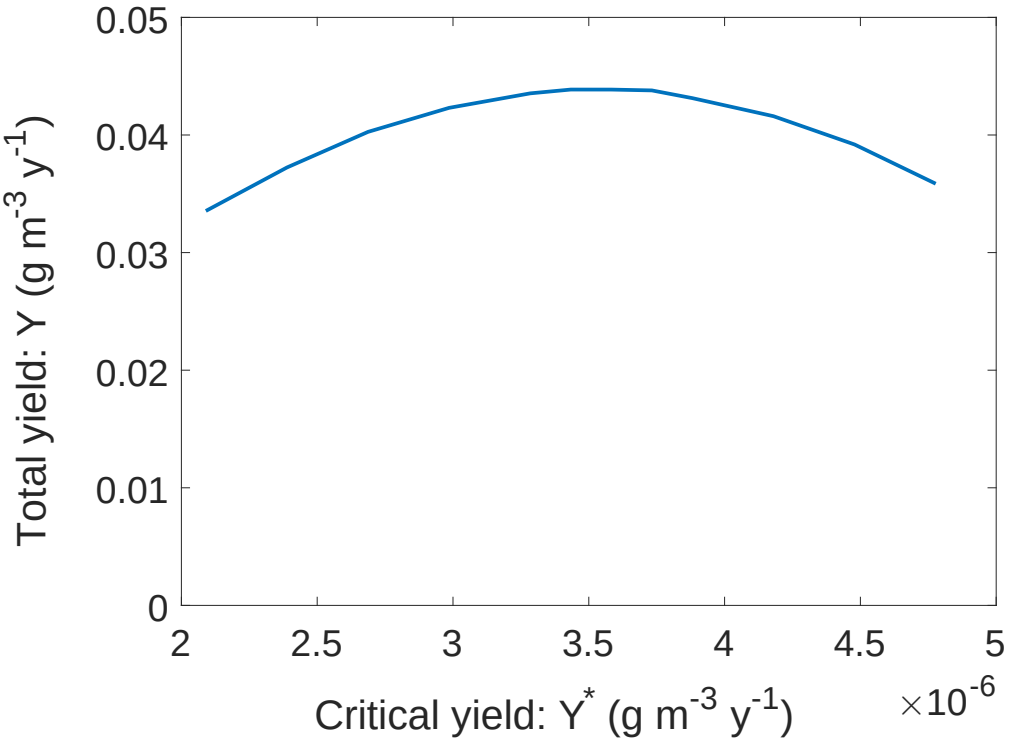


Figure 2.5: There is a peak in the total equilibrium yield. For a critical yield of approximately $3.5 \times 10^{-6}\mathrm{g\,m^{-3}y^{-1}}$ per fisher, a maximum equilibrium yield is obtained. Increasing or decreasing the critical yield from this level will result in a lower sustained yield over time.

2.5 Discussion

Size-based regulation methods of fisheries management are commonly used because of both economic and ecological concerns (Beddington et al., 2007). However, recent research suggests that alternative approaches to fisheries management may be better suited to preserving marine resources through minimising the impact of exploitation on those ecosystems (Garcia et al., 2012). Balanced harvesting, in which fish of a certain size are caught in proportion to their productivity at that size, has been suggested as an approach that will minimise disruption to fish populations and so will allow exploitation to be sustainable in the long term. For instance, balanced harvesting will reduce the number of BOFFFs being caught, and therefore increase the likelihood of continued reproductive success by the population. Furthermore, the economic justification that larger fish are more valuable on a price per unit weight basis in large-scale commercial fisheries, does not necessarily apply in the small scale artisanal fisheries. In such regions nutritional requirements, that small fish are equally able to meet, are a more significant concern. In these regions, fishing also provides a major source of employment, and so limiting the amount of targetable fish could also have an economic impact (FAO, 2014). Consequently, allowing harvesters in a small scale fishery to act without minimum size regulations could be worthwhile approach.

In each of the individual based models of fishing behaviour, fishers are seen to self organise over time into a Nash equilibrium solution. As the biomass spectrum has constant value over the entire range of targeted fish sizes, individual fishers will be unable to increase their own yield by changing their size-selectivity behaviour, no matter what their current target size is. This is as expected, as if fishers were able to change their own size-selectivity in order to increase their own yield, they would do so. Consequently any equilibrium solution must have all fishers obtaining the same yield. This result occurs for any choice of critical yield, with the long term solution in each case being that of equal yields for all fishers. The width of the flattened part of the biomass spectrum depends on the width of the targeted range of fish sizes. When there are more fishers in the fishery, the overall fishing pressure

increases, meaning there is an increase in the range of targeted fish sizes. This corresponds to an increase in the width of the flattened part of the exploited biomass spectrum, such that the deviation of the population size-spectrum from the unexploited level increases.

When the cost per unit effort of fishing is high, or the price of fish is low, a high critical yield will be required to keep fishers active in the fishery. High critical yields correspond to relatively low levels of exploitation, where the total size-selectivity behaviour is targeted at body sizes at which the population has high biomass. The large peak in the total size-selective fishing pressure ($F(w)$) occurs at the fish body size at which incoming biomass is used predominantly for reproduction rather than somatic growth. There is an accumulation of a large number of fish at and near this body size at any time, as they allocate their efforts towards reproducing, causing the local maximum in the biomass spectrum. As productivity is high here, exploitation must be significant in order for a stable equilibrium solution to occur. For higher levels of overall exploitation, with a constant biomass spectrum, fishers are able to obtain sufficiently large yields by targeting small fish, in particular fish smaller than the size at which the biomass peak occurs. This means an increased number of juvenile fish will not reach the size at which incoming biomass is allocated to reproduction, because the increased harvesting pressure is acting as an additional source of mortality for small fish. In this case there is not such an accumulation of fish of this size, and so the productivity of a population in this situation is not as high at this size. Consequently, a concentration of fishing pressure at that specific size range will not occur at equilibrium. Furthermore, the total size-selectivity is approximately proportional to the productivity of the population over the entire range of targeted fish sizes. Under high levels of exploitation, balanced harvesting is occurring, as an emergent result of the self-organisation of fishers into a Nash equilibrium.

The introduction of dynamics in the number of fishers extended the results from the individual based model with a fixed number of fishers, by showing that when the economic concerns of fishing are taken into account, a Nash equilibrium will still emerge. A new pa-

rameter based on costs that controlled the yield required to keep a single fisher active in the fishery was incorporated into the model. This in turn controlled the equilibrium number of fishers able to sustainably exploit the population, and thus the total fishing mortality at equilibrium. The fact that fishers were able to leave the system if their personal yields were low did not prevent remaining fishers eventually self-organising into a stable Nash equilibrium solution. Similarly, under high exploitation the equilibrium aggregate size-selectivity for remaining fishers was proportional to the productivity of the fish species, over the entire range of targeted body sizes. Consequently, a balanced harvesting situation still emerged from the model in the case where fishers' costs and revenue were taken into account. As such, the new dynamic individual based model produced results that were qualitatively similar to the existing individual based model. Further possible adaptations to the model which could be considered including allowing non-Gaussian size-selectivity functions in order to model other harvesting methods such as trawl fishing, and considering cooperative behaviour between fishers.

Chapter 3

The prisoner's dilemma structure of a fishery at bioeconomic equilibrium

3.1 Introduction

In this chapter, we will use game theory to model the choices about fishing effort, in a scenario where two fishers have the opportunity and ability to harvest a fish population. We will first develop a model for monopolistic exploitation by a single fisher, in order to determine the overall profit maximising behaviour in the fishery, which will then be used in the two-player game. Fishing effort measures the amount of fishing that can take place, and takes into account the time spent fishing and the availability of capital (for instance, in terms of boats or nets). Games provide a way to understand the strategic interaction between two or more individuals, or players. (Tadelis, 2013). In a game, each player makes a decision about which action to take. In doing so, they take into account their preferences for certain outcomes (usually in terms of the payoff they personally receive), as well as their own beliefs about how their rivals will act. Games therefore provide a natural way to model mathematically the decisions made by fishers, who independently wish to maximise their own profit from fishing, while taking into account how any other fishers' actions will impact fish stocks. In this chapter we will develop a non-cooperative game representing the decision

about fishing effort, by combining fish population models with models of the economic gains available from fishing.

The strategy a player chooses describes which action they will take at any stage of the game (Tadelis, 2013). The combination of all players' strategies is a strategy profile, and specifies all actions that will be taken during the entire play of the game. Based on the exact combination of actions taken at a given decision point in the game, each player will receive a payoff. In the context of our fisheries game, the payoff each player obtains will be the profit they obtain by fishing.

There are different solution concepts through which games can be understood. One of the most important solution concepts is the Nash equilibrium (Tadelis, 2013). In a Nash equilibrium, all players' strategies are best responses to each other, and no player can unilaterally deviate from this strategy in order to increase their own payoff (Tadelis, 2013). A contrasting concept is Pareto efficiency. An outcome is Pareto efficient if no player can be made better off by changing their strategy, without making another player worse off.

The well-known prisoner's dilemma (Rapoport and Chammah, 1965), which has payoffs satisfying $Su < Pu < Re < Te$ can be represented by:

		Player 1	
		Cooperate (C)	Defect (D)
Player 2	Cooperate (C)	(Re, Re)	(Te, Su)
	Defect (D)	(Su, Te)	(Pu, Pu)

Table 3.1: The standard payoff matrix for a two-player prisoner's dilemma. Entries are payoff vectors, where the first entry is the payoff to player 1, and the second entry the payoff to player 2.

A single instance of a game can be referred to as a one-shot game or a stage-game. By successively repeating an identical prisoner's dilemma stage-game between the same players,

who have a choice between the same actions each time, a multi-stage iterated prisoner's dilemma (IPD) can be developed (Tadelis, 2013). IPDs require another constraint on payoff values (Rapoport and Chammah, 1965), that is, $2Re > Su+Te$, in order to prevent collusion where players alternate between the strategy profiles (C, D) and (D, C). IPDs allow players to change their strategies at each iteration depending on both the actions of the other player and their own actions in the past. Simple strategy profiles, in which choices are based on only the last round of the game, are called memory-one strategies (Kraines and Kraines, 2000). Important pure memory-one strategies are

- $ALLC$: always cooperate
- $ALLD$: always defect
- TFT: tit-for-tat, cooperate on the first round and then match the other players previous action
- $STFT$: suspicious tit-for-tat, defect on the first round and then match the other players previous action

The strategies used by a population of a large number of players can be referred to as a population of strategies. An evolutionarily stable strategy (ESS) is one in which, when predominantly present in a population of strategies, is unable to be invaded by a differing "mutant" strategy that would have higher reproductive fitness (Smith and Price, 1973), in terms of better expected payoffs. Formally a strategy s is an ESS in a set of possible strategies S if it satisfies the following conditions (Smith, 1982)

- When playing against another player using s, a player using s must outperform a player using any other strategy $s' \in S, s' \neq s$.
- If a new strategy s' evolves in the population that performs equally well against the strategy s, then a player using s must perform better against a player using s' than a player using s' does.

No pure strategy can be evolutionarily stable in an IPD (Boyd and Lorberbaum, 1987), meaning that any population of strategies that are successful in terms of average payoffs can still be defeated by some other strategy or combination of strategies. Axelrod (1987) showed that in simulated tournaments between players in an IPD the TFT strategy resulted in the highest average score over all tournaments. If mixed strategies are allowed then a Pavlov (Win-Stay Lose-Switch or $WSLS$) will outperform TFT in similar simulations (Nowak et al., 1993). Another mixed strategy of note is generous tit-for-tat ($GTFT$), which behaves the same as TFT except that it has a small probability of forgiving an opponent who defects, by choosing to still cooperate in the next round. This allows a situation of recurring mutual punishment (such as will occur with TFT vs. $STFT$) to be avoided.

Both pure and mixed memory-one strategies can be described in the form (p_1, p_2, p_3, p_4) which give the probability of cooperation by a player using that particular strategy, based on the outcome of the previous iteration of the game. Then p_1 is the probability of cooperating after cooperation by both players, p_2 after self cooperation and opponent defection, p_3 after self defection and opponent cooperation, and p_4 after defection by both players. We can then define, for example, $ALLC$ as $(1, 1, 1, 1)$, or TFT as $(1, 0, 1, 0)$. However, any strategy relying on only the immediately previous iteration can be described in this way.

In this chapter, we will show that in certain circumstances, the exploitation of a renewable resource by multiple fishers constitutes an iterated prisoner's dilemma with respect to each fisher's choice of harvesting rate.

Various different approaches have been used to model the growth of a population, based on differing views on whether populations are regulated by density dependent biotic effects, or density independent abiotic effects. Density dependent effects include things like intra-community competition and disease that will have a relatively larger impact on high-density communities than those of low-density. On the other hand, some simple population models

include density independent effects such as weather or climate fluctuations, which have a similar effect on communities regardless of population size, to be the predominant factor in regulating population growth. In reality, both aspects likely play a part in controlling population growth to some extent. We will focus on the effects of density-dependent controls on populations.

Game theory has been extensively applied to the study of fisheries, with contributions having been thoroughly reviewed by Sumaila (1997) and by Bailey et al. (2010). Gordon (1954) showed that the overexploitation of marine resources occurs because of their common property nature, meaning they yield no economic rent, and found the optimal level of harvesting for a single manager. Munro (1979) was the first to approach this problem using game theory, in the context of two countries both exploiting a marine resource that straddles the boundary of their Exclusive Economic Zones, and in particular how asymmetries in discount rates, costs, and consumer tastes affect their outcomes. He combined the Schaefer (1954) (also known as the Gordon-Schaefer) model of fisheries dynamics with an objective function giving the discounted net cash flow from the fishery. It was found that allowing side payments from one player with low costs to the player with high costs reduced the game to a sole manager situation, and thus allowed for optimal harvesting.

Levhari and Mirman (1980) contextualised their approach as a 'great fish war' for cod between Iceland and the United Kingdom. Their discrete time model particularly considered population dynamics, and the fact that the density of the underlying marine resource changes as a result of the actions of both players, meaning that each player must consider the effect of their opponents actions of stock levels. This created what they call a 'dynamic externality' (Levhari and Mirman, 1980) in the interaction between the two players. Dockner et al. (1989) also considered a nonzero-sum game between two players exploiting a fishery with an underlying Gordon-Schaefer model. Their model was duopolistic rather than one in which there is sole ownership of the fishery, or open access to it. An important distinction is that

the price of landed fish depended on the total quantity that is harvested by both players, such that they are in a oligopolistic rather than competitive market. They considered both Nash games, in which players choose their actions simultaneously, and Stackelberg games (Von Stackelberg, 2010), in which a market leader first makes a decision, and the other player responds.

A bioeconomic model of the exploitation occurring in a fishery can be created by coupling a growth model of a fish population with a behavioural model of the actions taken in exploiting that population. This model can be formulated as a non-cooperative game, where one or more fishers chooses an action (their own fishing effort) such that an outcome occurs (their own yield they have harvested from the fishery) for which they receive a payoff (the profit obtained from fishing). If the resultant game has a form for which results in game theory are well known, such as the prisoner's dilemma, then these results can be extended to the underlying fishing model, giving insight into which types of bioeconomic situations will allow socially optimal outcomes to occur.

3.2 Population models

We will now give a brief description of the biological population models that will be used in this chapter. One simple model of density dependent population growth is the logistic equation, also known as the Pearl-Verhulst equation, in which the relevant population has a per capita growth rate that decreases linearly with population density. Under such a model, the population has a growth rate parameter r and a carrying capacity K which is the stable equilibrium solution of the model. There is also another equilibrium solution where the population is fully depleted and is therefore unable to be replenished. The growth of the population u is given by:

$$\frac{du}{dt} = f(u), \quad f(0) = f(K) = 0 \tag{3.1}$$

for a continuous time model, or:

$$u_{n+1} = f(u_n) \quad f(0) = 0,\ f(K) = K \tag{3.2}$$

for a discrete time model. Some common choices for the growth function include:

- The logistic equation

$$f(u) = \frac{du}{dt} = ru\left(1 - \frac{u}{K}\right) \tag{3.3}$$

- The Gompertz growth equation

$$f(u) = \frac{du}{dt} = \alpha u \log\left(\frac{K}{u}\right) \tag{3.4}$$

- The Beverton-Holt map

$$f(u) = u_{n+1} = \frac{Ku_n}{u_n + (K - u_n)e^{-rT})} \tag{3.5}$$

Such models of population growth can be simply adapted to become models of harvesting by adding a component of mortality due to resource extraction. We assume that the rate of fishing mortality is proportional to the population density at any given time. With the addition of a harvesting term the behaviour of the dynamical system representing the model changes. The general form for a continuous time model is then:

$$\frac{du}{dt} = f(u) - Fu \tag{3.6}$$

Where F is the constant rate of fishing effort. For example, the logistic growth model can be extended to become the Gordon-Schaefer model of fishing mortality by assuming that the rate of fishing mortality is proportional to the population density at any given time:

$$\frac{du}{dt} = ru\left(1 - \frac{u}{K}\right) - Fu \tag{3.7}$$

As with the logistic equation there are two equilibria. One is at

$$u^* = K\left(1 - \frac{F}{r}\right) \tag{3.8}$$

representing the long-term population level under fishing effort F. The other occurs again at $u^* = 0$, representing the case where the population is fished to extinction. At the parameter values $F = r$ a transcritical bifurcation occurs, in which there is a qualitative change in the behaviour of the system. This means that for $F < r$, the equilibrium point $u^* = K(1 - \frac{F}{r})$ is stable and the point $u^* = 0$ is unstable, while for $F > e$ these points switch stability. In real terms this means that when the fishing mortality is less than the intrinsic growth rate of the population, a "sustainable yield" (or equilibrium harvest rate) will be possible. However when the fishing mortality is greater than the population growth rate, the population will eventually be fished to extinction.

In a similar fashion, the Gompertz growth equation can be extended to the Fox surplus yield model

$$\frac{du}{dt} = \alpha u \log\left(\frac{K}{u}\right) - Fu. \tag{3.9}$$

The Fox model also has two equilibria, at $u^* = 0$ and at $u^* = K\exp(\frac{-F}{r})$. Here a bifurcation occurs at the point $\frac{F}{r} = 0$.

3.3 Modelling framework

Let there be a fish population of density u that exhibits density dependent growth according to the differential equation

$$\frac{du}{dt} = f(u) - Fu, \quad f(0) = f(K) = 0 \tag{3.10}$$

where r is the population growth rate, K is the carrying capacity, and F is the exploitation rate. Without loss of generality, we can let $K = 1$. This population has no size-structure, with u simply representing the biomass density of the population as a proportion of its carrying capacity.

We begin by considering a situation in which a single fisher has the sole ability (whether through access or property rights) to exploit the fish population. We assume a fisher will harvest at constant effort F during a fishing period of length T. The profit obtained by the fisher during a single period is

$$P = BF \int_0^T u(t)dt - cFT \tag{3.11}$$

B is the price per unit biomass of fish, and c is the cost per unit effort. We will then extend the model to allow the case of two fishers exploiting the same population. Each fisher will have independent decision making behaviour, and so each fisher will be able to select their own fishing effort F_1 (or F_2) at the beginning of each time period, in order to satisfy their own optimality condition. For multiple fishers, the optimality condition is to maximise that fisher's own profit, and not the total payoff obtained by exploiting the fishery. We will then further extend the model to an iterated game, in which the single stage game between the two fishers is repeated, with the state of the population, and each player's action, being able to inform the players' future decision making behaviour. In the iterated game, each fisher has the ability to change their current effort $F(n)$ at the beginning of the nth fishing period.

We will consider three implementations of this model in this chapter. These are

- **Fast dynamics**: This is a limiting case in which the population has a high growth rate and so reaches biomass equilibrium quickly compared to the time-scale on which changes in fishing effort (F) occurs. In this case we are able to derive analytically results including payoffs to individual fishers, and population biomass levels.

- **Slow dynamics**: This is a numerical model in which the population has sufficiently slow dynamics that growth and harvesting occur on similar time-scales.

- **Separation of time-scales** A numerical model of a fish population with a Beverton-Holt map as the stock recruitment relationship. In this case reproduction occurs at discrete intervals (at the beginning of each period), representing the regularly occurring seasonal spawning process of the fish population. Exploitation happens continuously throughout the period, and as such occurs in this case over a much longer time-scale than population replenishment.

Figure 3.1 is a schematic diagram depicting the key differences between each of these models, in particular with their approach to time-scales. Each implementation considers discrete fishing periods of equal length T. At the beginning of each fishing period, each fisher individually chooses a level of fishing effort, which remains constant during this time. Figure 3.1 (a) shows an example of fishing efforts over 5 time periods, with these levels of effort being applied to each of our harvesting models in $(b) - (d)$. Harvesting occurs continuously throughout each period in all of our models. In the equilibrium population model in Figure 3.1 (b), population growth is assumed to instantaneously reach its new equilibrium level under the active fishing effort, at the beginning of the period. Population growth occurs continuously during this time such that biomass levels are constant throughout each single fishing period. In (c) we see the logistic growth model with slow dynamics. Here resource growth also occurs continuously according to its governing differential equation. However, as equilibrium is not quickly reached, population biomass levels will change over a fishing period. Finally, (d) shows the response of the Beverton-Holt recruitment process to fishing. In this case, population growth occurs discretely at the beginning of each period, with no growth occurring during it. Therefore any population change at this time is entirely due to the effect of exploitation.

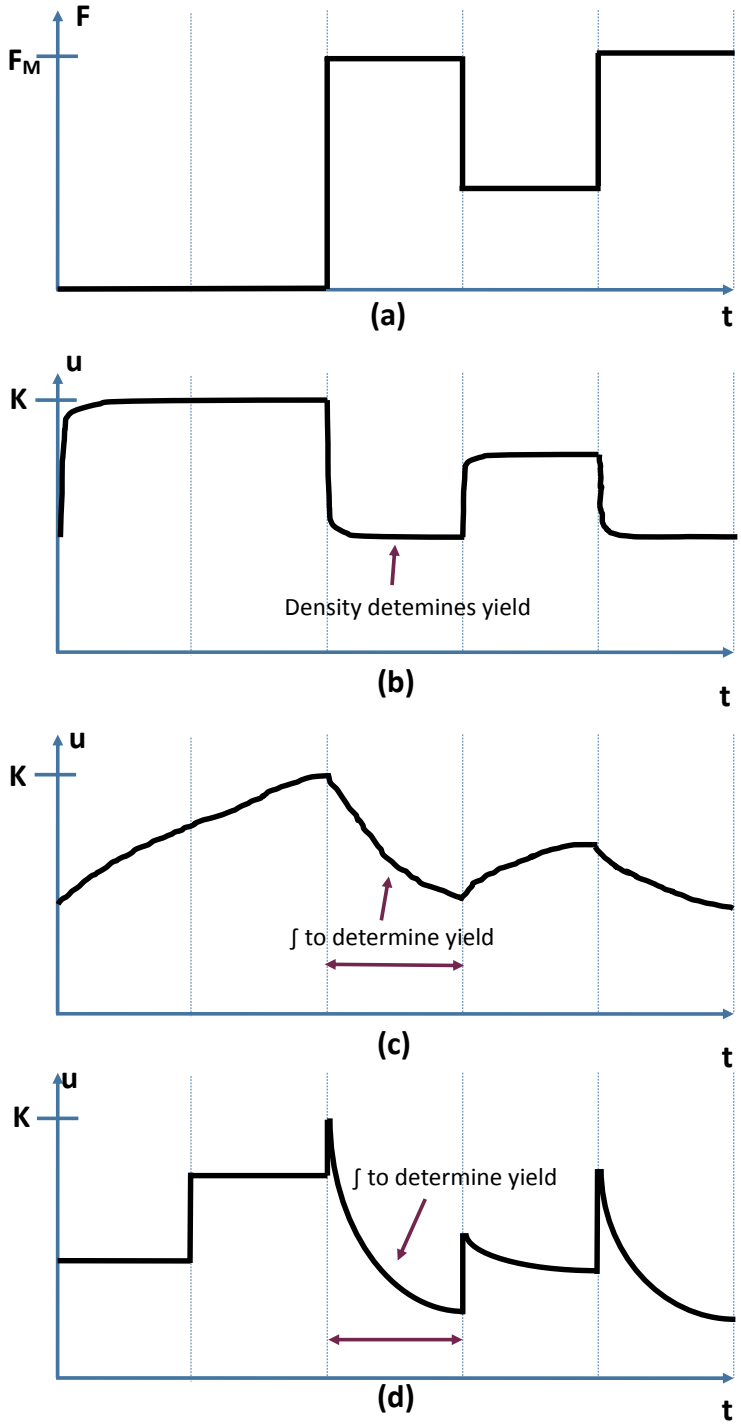


Figure 3.1: Schematic diagram depicting the differences between how each of our models respond to the same level of exploitation, with the same initial conditions. (a) shows the level of fishing effort F being applied to a population over time t. (b) shows how the population biomass u responds when it has logistic growth with fast dynamics. (c) is the biomass response to this fishing effort for logistic growth with slow dynamics. (d) is the biomass level for a Beverton-Holt growth model. In each case, exploitation occurs continuously throughout each time period, with yield calculated by integrating over this time. The models differ in how each deals with the natural growth of the population over time.

3.4 Fast dynamics

3.4.1 Equilibrium assumption

We begin by assuming that $1/r \lll T$, so that the population will instantaneously reach equilibrium, that is, the population density is in its equilibrium state (for the applied level of fishing effort F) for the entirety of the fishing period. Any transient behaviour from the state of the population before exploitation reaching equilibrium is assumed to be insignificant and ignored. This equilibrium population density is $u^*(F)$ where

$$f(u^*(F)) - Fu^*(F) = 0. \tag{3.12}$$

In this model, we consider a population that experiences logistic growth, and so is an example of the Gordon-Schaefer model of fisheries. Population dynamics are given by

$$\frac{du}{dt} = ru\left(1 - \frac{u}{K}\right) - Fu, \tag{3.13}$$

which means the population has an equilibrium density

$$u^* = K\left(1 - \frac{F}{r}\right). \tag{3.14}$$

3.4.2 Monopolist solution

We first consider the situation in which there is a single fisher who has the exclusive and unique ability to exploit a fishery. We do this to derive the monopolist profit level, which is the greater economic return that can be obtained from the fishery and is the profit corresponding to the MEY. This monopolist profit can then be used to determine a total payoff from the fishery that is Pareto efficient, in a situation in which there is ore than one fisher. Under our equilibrium assumption, $u^*(F)$ is clearly not changing over the fishing period, and

so the equilibrium profit obtained by the fisher will be

$$P^*(F) = BFu^*(F)T - cFT \tag{3.15}$$

We assume that the sole fisher is motivated by profit, and not other considerations such as conservation. They will therefore act to maximise the profit they are able to obtain from the fishery. We let F^M be the *monopolist effort*, the effort level which maximises the profit that can be extracted from the fishery, chosen so that

$$\frac{dP^*}{dF}(F^M) = 0 \tag{3.16}$$

where $F^M > 0$. In the Gordon-Schaefer model, the monopolist profit is

$$P^* = BF^M K\left(1 - \frac{F^M}{r}\right)T - cF^M T \tag{3.17}$$

Without loss of generality we can let the length of a fishing period be $T = 1$, and recalling that we are also able to let $K = 1$. Then the profit maximising level of harvesting is

$$F^M = \frac{r}{2}\left(1 - \frac{c}{B}\right) \tag{3.18}$$

3.4.3 One-stage two-player game

What if there are two fishers who have identical access to the fishery? We assume they receive the same price for fish, experience the same costs of fishing, and have the same profit motivations driving their decision making behaviours. In this chapter, we will use $i = 1, 2$ to represent a player, and $j = 1, 2$ to represent the other player. We therefore let F_i be the effort level chosen by the ith fisher. The total fishing effort is $F = F_i + F_j$, and so we note that the equilibrium population density is dependent on the decisions made by each player about their own effort. Because profit is dependent on biomass, this means that P_i, the profit received by the ith player depends on both their own effort, and the other player's

effort.

We wish to know what the ideal behaviour is for each player under these conditions. The socially optimal result (a Pareto equilibrium) will occur when the total payoff (that is the economic profit) from the fishery is maximised, which will simply happen when the total fishing effort of both fishers is equal to the monopolist effort, i.e. $F^M = F_i + F_j$. Because of the symmetry of the fishers' situation, we can consider a case where they equally share these profits, meaning $F_i = F_j$ and so $P_i = P_j$. Consequently we define a "Pareto equilibrium" (or "cooperative") level of fishing effort $F^C = F^M/2$ which will allow this to happen when both players choose this action. Note that while any combination of efforts which sums to the monopolist effort will result in a Pareto optimal solution, we are restricting our consideration to the particular case when that total effort is distributed equally between the fishers, in order to define an action analogous to cooperating in the traditional prisoner's dilemma. In our model with Gordon-Schaefer population dynamics, the Pareto effort is

$$F^C = \frac{r}{4}\left(1 - \frac{c}{B}\right) \tag{3.19}$$

Analogously with the prisoner's dilemma, we will refer to the choice by a fisher to fish at the Pareto effort F^C as cooperating, as like the prisoner's dilemma when both players choose this action a Pareto optimal solution will result. Each decision maker, however, is not motivated by the goal of reaching the best social outcome, but is instead trying to maximise their own profit, by playing a strategy against the other player. They are only concerned with their competitor's behaviour insofar as it affects their own outcomes. The players have identical preferences for profit, and furthermore each knows this; this is a game of complete information. We consider, analogously to Equation (3.15), the profit obtained by Player i:

$$P_i = BF_i u^*(F_i, F_j)T - cF_iT \tag{3.20}$$

which is maximised when

$$\frac{\partial}{\partial F_i} P_i\left(F_i, F_j\right) = 0 \tag{3.21}$$

giving our optimality condition as

$$\frac{\partial P_i}{\partial F_i} = B - \frac{2B}{r} F_i - \frac{B}{r} F_j - c = 0 \tag{3.22}$$

Noting the dependence of this optimality condition on both players' efforts, we denote Player is best response $s_i^*(F_j)$ as a function of Player j's effort; for each possible choice of effort by Player j this is Player i's "best response" strategy. Noting the symmetry of the players' situations, we also have Player j attempting to maximise their profit with respect to the behaviour of Player i, and so they also have a best response strategy $F_j = s_j^*(F_i)$ to Player i's actions. Because of symmetry, these best responses must be the same,

$$s_j^*(F) = s_i^*(F) \tag{3.23}$$

We consider the case where both players choose to play best responses to the other player's actions, which by definition means there is a Nash equilibrium at (F_i, F_j). The actions each must therefore be the same, that is $F_i = F_j$, and so Equation (3.22) is easily solvable. We let F^D, the "Nash effort" be this effort value where both players are playing mutual best responses to the other's strategy. The Nash effort in our model corresponds to the choice to defect in a traditional prisoner's dilemma.

Solving Equation (3.22), we get our best response function

$$s_i(F_j) = \frac{r}{2} - \frac{F_j}{2} - \frac{cr}{2B} \tag{3.24}$$

which gives the value of F_i which will maximise Player i's profit for a given value of F_j. This is always a maximum since

$$\frac{\partial^2 P_i}{\partial F_i^2} = \frac{-2B}{r} < 0, \tag{3.25}$$

as B and r are positive. We also see that

$$\frac{\partial P_i}{\partial F_j} = -\frac{B}{r} F_i \tag{3.26}$$

meaning that when Player i is fishing at any non-zero effort level, their own profit will decrease as a result of any increase in Player j's fishing effort. This is the Tragedy of the Commons (Hardin, 1968), where the impacts of the decreased biomass caused by Player j's increase in fishing effort are shared between all players, while Player j takes all the reward due to their increased yield.

From Equation (3.24) and the symmetry of the fishers behaviour, we then obtain the "Nash effort"

$$F^D = \frac{r}{3}\left(1 - \frac{c}{B}\right). \tag{3.27}$$

We will refer to the choice of effort F^D as defecting, analogously to the prisoner's dilemma example, as fishing at F^C represents a choice by a fisher to maximise their own profit, with respect to whichever action the other fisher may take.

While a continuous range of potential fishing efforts is allowable in a single-stage game, in order to form meaningful conclusions about the fishers' decision making behaviour we can restrict these actions to create a sub-game in which the results of our model hold true. By limiting the fishing effort which may be chosen by each player down to a possible two discrete choices, we can therefore consider the model in the form of a two-player game with two possible (and identical) actions for each player, similar to the structure of many common games. These actions are to "cooperate", or to fish with effort F^C, or to "defect" and fish

	Reward	Sucker	Temptation	Punishment
F_i	$\frac{\alpha}{4}$	$\frac{\alpha}{4}$	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$
F_j	$\frac{\alpha}{4}$	$\frac{\alpha}{3}$	$\frac{\alpha}{4}$	$\frac{\alpha}{3}$
u^*	$\frac{1}{2}\left(1+\frac{C}{B}\right)$	$\frac{1}{12}\left(11+\frac{7C}{B}\right)$	$\frac{1}{12}\left(11+\frac{7C}{B}\right)$	$\frac{1}{3}\left(1+\frac{2C}{B}\right)$
P_i	$\frac{\beta}{8}$	$\frac{5\beta}{48}$	$\frac{5\beta}{36}$	$\frac{\beta}{9}$
P_j	$\frac{\beta}{8}$	$\frac{5\beta}{36}$	$\frac{5\beta}{48}$	$\frac{\beta}{9}$

Table 3.2: A summary of the efforts (actions) of each player, their profits (payoffs), and the equilibrium population density for each possible outcome of the game. We define the parameters $\alpha = r\left(1-\frac{C}{B}\right)$ and $\beta = \frac{r}{B}\left(B-C\right)^2$

with effort F^D, as derived in this section. It is clear from Equations (3.27) and (3.19) that $F^C < F^D$, such that defecting means exploiting the fish population at a higher level than cooperating. Recall from Subsection 1.1.2 that higher fishing efforts result in higher yields, but not necessarily higher profits. In this chapter we therefore expect the yields obtained when players defect to be higher than when they cooperate, and consequently that the resultant population levels will be lower under defection.

This game can be formulated as a single-stage game which will in the next section be repeated to form an iterated game. As we have analytical equations for the two possible actions in the game (for any parameter values), we can obtain the equilibrium population densities and the players' rewards for each possible outcome. In such a two-player game, there are four possible rewards:

- "Reward" (Re): Both Players cooperate
- "Sucker" (Su): Player i cooperates and Player j defects
- "Temptation" (Te): Player i defects and Player i cooperates
- "Punishment" (Pu): Both Players defect

The results for each of the payoffs (for Player i) are summarised in Table 3.2. Clearly the

	Reward	Sucker	Temptation	Punishment
F_i	0.3130	0.3130	0.4789	0.4789
F_j	0.3130	0.4789	0.3130	0.4789
u^*	0.5347	0.4530	0.4530	0.3838
P_i	0.1048	0.0792	0.1212	0.0880
P_j	0.1048	0.1212	0.0792	0.0880

Table 3.3: A summary of the efforts (i.e. action) of each player, their profits (payoffs), and the equilibrium population density for each possible outcome of the game when the population grows according to the Fox model. Here $r = 1$, $K = 1$, $B = 1$ and $C = 0.2$.

payoff ordering

$$Su > Re > Pu > Te \tag{3.28}$$

means that our restricted game is equivalent to a standard prisoner's dilemma. Consequently, as expected, in a one-shot game the solution is a Nash equilibrium in which both players defect, and existing results apply with regards to the success of strategies in the repeated game.

We also considered the Fox model of fishing, an alternative continuous time population model for which numerical results were used to obtain the possible effort choices for each player. This model did not produce qualitatively different results, with a prisoner's dilemma situation emerging as above. Table 3.3 gives numerical values for efforts, payoffs, and biomass for each possible outcome of the Fox model for a specific set of parameters. The relationship between payoff values for each outcome was invariant to parameter changes, meaning the 2-player game was equivalent to prisoner's dilemma for all parameter values.

3.4.4 Repeated two-player game

We can form an iterated game by repeating our single shot game over an infinite time horizon (with discounting of payoffs by a factor of δ). Note that because the single shot game was a prisoner's dilemma when restricted to 2 actions, and because of the clear ordering of outcomes $2Re > Su + Te$ seen in Table 3.2, this repeated game will be consistent with an IPD. Each

player knows the state of the fishery (that is its biomass u) at each point in time, representing the ability to perfectly assess stock levels. They are able to infer directly the moves their opponent makes in each stage-game, as they know both their own action at that time and the population dynamics. This iterated game is therefore a game of perfect information. At the outset of the repeated game, players choose a strategy profile defining the action they will take given the state of the game in previous iterations. The list of considered strategies here (or in similar cases throughout the chapter) are:

- Always cooperate ($ALLC$) $(1, 1, 1, 1)$
- Always defect ($ALLD$) $(0, 0, 0, 0)$
- Tit-for-tat (TFT) $(1, 0, 1, 0)$, cooperate in the first stage-game
- Suspicious tit-for-tat ($STFT$) $(1, 0, 1, 0)$, same as tit-for-tat, but defect in the first stage-game
- Generous tit-for-tat ($GTFT$) $(1, 0.01, 1, 0.01)$, has a 1% chance of forgiving defection by the other player
- Very generous tit-for-tat ($VGTFT$) $(1, 0.1, 1, 0.1)$, has a 10% chance of forgiving defection by the other player
- Pavlov or win-stay-lose-shift strategy ($WSLS$) $(1, 0, 0, 1)$
- Suspicious win-stay-lose-shift ($SWSLS$) $(1, 0, 0, 1)$ same as Pavlov, but defect in the first stage-game
- Sly $(0, 0, 1, 1)$ defect if there was a favourable payoff in the last stage-game

In Figure 3.2 we see the profit obtained by Player 1 in this equilibrium density IPD, when both players choose different combinations of these strategies. In this heatmap, rows represent the strategy employed by Player 1, with columns representing the strategy played by their opponent. In each case in Figure 3.2 we see that the relative performance of strategy

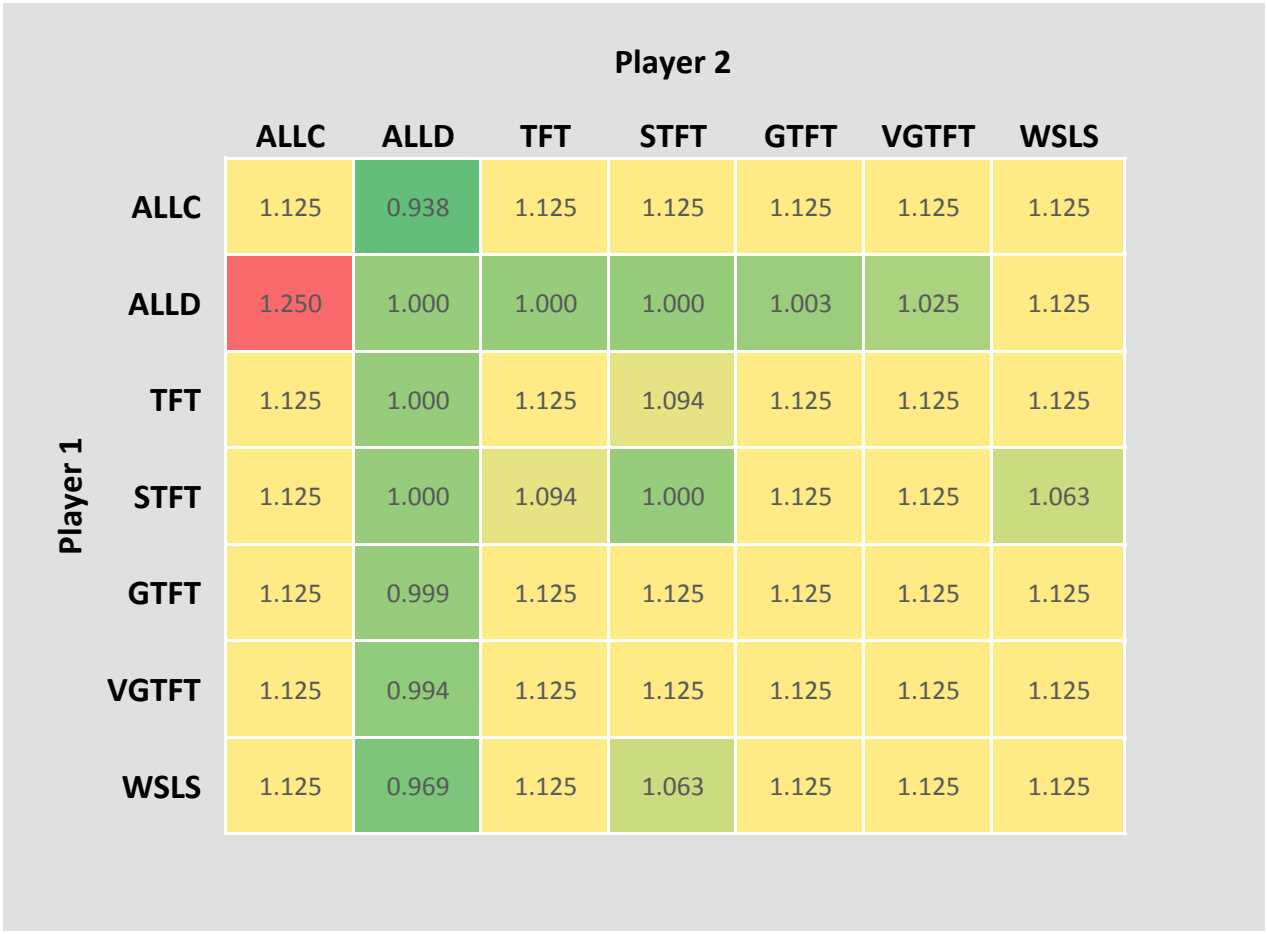


Figure 3.2: A heatmap for the average profit obtained by Player 1 in an IPD, when different strategy profiles are played. At each iteration of the game, population equilibrium is reached, as the population has fast resource dynamics. Parameter values are $r = 20$, $K = 1$, $B = 1$, and $c = 0.2$. The payoffs are normalised by the value of the Nash equilibrium (the ALLD vs ALLD payoff) which here is 1.422. Relative performance of strategy profiles are here as expected from the standard results for a prisoners' dilemma.

profiles is as expected from existing results for an infinite time horizon IPD (Axelrod, 1981; Imhof et al., 2005). For example, when $ALLC$ plays $ALLD$ it is dominated, as the $ALLD$ player will receive the temptation payoff in each stage-game, while the $ALLC$ player only obtains the sucker payoff. In an evolutionary game theory scenario, a population of $ALLC$ players is therefore vulnerable to being invaded by $ALLD$ players(Imhof et al., 2005). In a series of well-known simulations comparing many strategies for an IPD, tit-for-tat was the most successful strategy (Axelrod, 980a,b) because it is "nice", meaning that a TFT player will never defect first. It is provocable into retaliation if there are deviations from cooperation, and quickly forgiving, as if a player that had defected begins to cooperate the TFT player will resume cooperation (Axelrod, 1981). Looking at Figure 3.2, we see that TFT performs relatively well against each strategy. In particular, when TFT plays $ALLD$, they receive the same payoff (the Nash equilibrium). TFT is resilient to being taken advantage of by $ALLD$. Furthermore, when TFT plays itself, it receives the cooperative reward payoff, which is greater than the punishment payoff. No strategy performs better against TFT than TFT performs against that other strategy, and no strategy performs better against TFT than TFT performs against TFT. TFT is therefore robust to being invaded by $ALLD$, as $ALLD$ does not strictly outperform it. Many of our other considered strategies are variations of TFT. The type of variation is responsible for differing relative performance against other strategies. For example, a forgiving strategy such as $GTFT$, which has a small chance of cooperating after a defection, can break the cycle of alternating cooperation and defection that happens when a TFT variant plays $STFT$. The downside to the generosity is that the $GTFT$ is more susceptible than TFT to being taken advantage of by the uncooperative $ALLD$.

3.5 Slow dynamics

The following section approach is an attempt to deal with populations in which the population growth rate r is low, or changes in fishers' fishing effort F occur more frequently, such that the fish population does not quickly reach equilibrium during any single fishing period. This corresponds to the situation depicted in Figure 3.1 (c), where reproduction and exploitation happen on a similar time-scale. In this model, players cannot directly observe the state of the fishery; this is a game of imperfect information. This represents a situation where fishers do not have accurate estimates of stock levels, and of population growth parameters. As in the previous model, we consider both the case where there is a single fisher with exclusive access, and the case where there are two identical fishers, each aiming to maximise their own profit. In the two player case, as before we will obtain two possible actions, a "Pareto" effort analogous to cooperating in a prisoner's dilemma, and a "Nash" effort which occurs when a player chooses to defect. The constructed two-player, two-action symmetrical stage-game will then be repeated in the final part of this section.

3.5.1 Monopolist solution

In this model, we assume a monopolist fisher has limited knowledge about the targeted fish population. The fisher knows the population has logistic growth, the carrying capacity, and population growth rate. They do not however know the actual amount of population biomass at any time, and are only able to infer it through consideration of their own yield. As before, the monopolist behaviour is investigated in order to determine the profit-maximising (MEY) action in the fishery, such that we can derive a Pareto efficient combination of fishing behaviours in the two-player situation.

The single player, as in the previous section, will act to maximise the payoff they receive by choosing the effort with which they exploit the fishery. How should they do this if they do not know the current population density? From the previous section (the fast dynamics

case), we know that the effort F^M (from equation (3.18)) will maximise the payoff obtained at equilibrium. With slow dynamics, the population will still eventually reach an equilibrium density (with constant exploitation). Therefore if the monopolist fisher has to set a constant strength of fishing effort, for the slow resource case we will have the same payoff maximising F^M as in the fast resource case. We note that because of the slower dynamics, the profit obtained by a monopolist over a period of fishing at effort F^M will likely be different than that at equilibrium.

3.5.2 One-stage two-player game

We now turn to the two-player, single-shot game case with slow resource dynamics. As before, each player will choose their action to maximise payoff, given their knowledge of the population behaviour, and their expectations of their opponent's behaviour. In a single-shot game, each player chooses an action once, with this combination of actions determining each player's payoff. As the players do not know the current state of the population, we can assume that they would act as if the population was at, or was able to reach, equilibrium. Consequently both players know the efforts F^C and F^D at which they should fish in order to cooperate or defect, as calculated in the fast resource dynamics case with logistic growth dynamics. They also have a concept of the profit and yield they should receive at equilibrium from each of the possible outcomes of this game. That is, they know the yield they would receive in the long term if fishing at the Pareto effort or the Nash effort, given the other player's behaviour. Because the population dynamics in this case can be slow, such that the fish population does not reach equilibrium quickly in the fishing period, the actual yield gained by a fisher will possibly be different than the expected equilibrium value.

3.5.3 Repeated two-player game

As before, the single stage-game with two players can be repeated to form an iterated game, in which each player will choose a strategy profile that allows them to change their fishing effort at the beginning of each iteration, based on the state of the game. Recall that

each fisher's effort will, however, remain constant throughout an single fishing period. In order to make the action space numerically feasible, we will assume that the efforts able to be chosen by each player are the Nash effort F^D and the Pareto effort F^C detailed previously.

How does having slow dynamics affect the iterated version of this game? The fishers do not have any ability to actually know the current population biomass at any given time, and therefore also do not know the previous actions of the other player. They must instead indirectly measure population biomass by considering their biomass yield Y during the fishing period, the actual amount of fish they catch:

$$Y = \int_0^T F_i u(t) dt \tag{3.29}$$

A players payoff is directly related to their own yield, as:

$$P_i = BY_i - cF_iT \tag{3.30}$$

Each fisher is also unable to observe directly the other player's action during any fishing period, and therefore can only attempt to discern it by considering their own yield. If the yield obtained by Player i is less than they expected during the nth fishing period, then it is reasonable for them to assume that Player j defected during that period. We therefore define a new parameter Y^*, the critical threshold yield, which is the minimum yield, when cooperating, that either player will accept in order to be convinced that the other player has also cooperated during that fishing period. We choose this threshold yield to be the midpoint between the yield obtained when both players cooperate at equilibrium, and the yield obtained when both players defect at equilibrium. This parameter allows us to develop imperfect information variants of the memory-one strategies describe in Section 3.4.4. These strategies behave similarly, except that decisions about opponent's prior behaviour come from indirect beliefs based on the threshold yield, rather than direct observation of actions.

For a sufficiently large number of fishing periods that the effects of transient behaviour can be disregarded, we use numerically determine the population dynamics and the yields obtained by each player. Each fishing period constitutes a single iteration of a repeated two-player game. Each player chooses a memory-one strategy to determine their choice of action at the beginning of each fishing period. As their opponent's action is not definitively known, these strategy profiles are not functions of the other player's actions, but of the yield obtained. Effectively, if the yield of Player i during the nth fishing period is less than that threshold yield:

$$Y_i(n, n+T) < Y^* \tag{3.31}$$

then Player i has the belief that Player j defected during the previous fishing period, and if Player i's yield is greater than the threshold yield:

$$Y_i(n, n+T) > Y^* \tag{3.32}$$

then Player i has the belief that Player j cooperated during the previous fishing period, allowing the state of the population, according to each player's respective strategy profiles, to be determined.

Using this method of determining each players' perception of their opponent's actions, we use simulations to numerically calculate the performance of different strategies against each other. For each possible combination of considered strategies, a large number of iterations of an IPD are run, with the population biomass, fishing efforts, and payoffs being determined over this time. Eulers method was used to solve (3.13) in order to determine the population biomass over time, given the present combination of fishing efforts. In each case the initial condition of the population was $u = K$, meaning the population was initially at its unexploited carrying capacity level. IPDs for each combination of strategies were simulated a large number of times, with average outcomes calculated. The normalised average profit obtained by Player 1 for each combination of strategies is given in Figure 3.3. We note the

normalisation factor for profit is much lower than in the high growth rate case (in part due to the dependence of profit on growth rate), meaning that the actual profits obtained in this case are lower than in the previous section. Furthermore, the relative performance of different strategies are not identical to those in the fast dynamics model. For example, in this iterated game, $ALLD$ outperforms TFT, receiving a higher payoff when those strategies are played against each other. This suggests that TFT is not robust to invasion by $ALLD$ here. When these two strategies play each other in an IPD, we expect continual mutual defections, with neither strategy outperforming the other in the long term. Here, it appears that the $ALLD$ player can take advantage of the TFT player; in some number of stage-games, the slow dynamics mean that the TFT receives a yield above the threshold level, and so they assume the $ALLD$ player cooperated in that previous stage-game. This causes the TFT player to get the sucker payoff while the $ALLD$ player gets the temptation payoff. Overall, the TFT player will receive some combination of punishment and sucker payoffs, and the $ALLD$ player a combination of punishment and temptation in the same ratio. In contrast to the fast dynamics model, TFT does not perform as well against any strategy as that strategy does against TFT. Variations of TFT ($STFT$, $GTFT$, $VGTFT$) perform identically to it, while the Pavlov ($WSLS$) strategy performs worse against $ALLD$ than TFT does. We conclude for this iterated game with slow dynamics, our imperfect information strategies do not result in the same payoff structure as with fast dynamics in an IPD.

3.6 Separation of time-scales

Our third and final approach is motivated by the biological nature of stock recruitment in fish populations, where population growth due to stock recruitment often occurs generationally in discrete spawning periods (Booth and Beretta, 1994; Allen and Barker, 1990; Campana, 1996). In addition to this, a major goal of this thesis is to understand the effect of population stochasticity on optimal fishing behaviour. It is therefore worthwhile to develop a model in which noise in the population growth process can be added in the future. This interim model

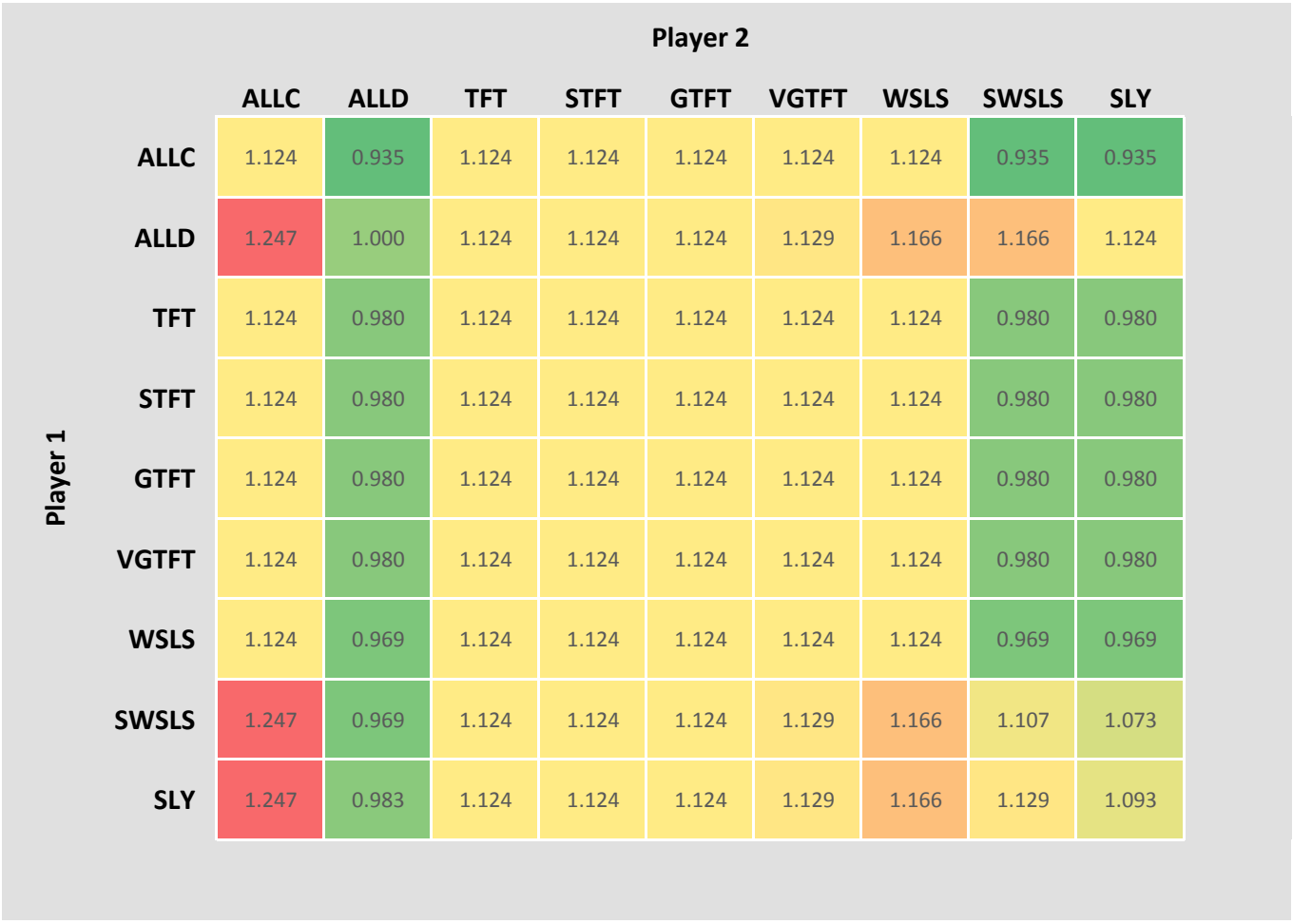


Figure 3.3: Payoff comparison heatmap for strategy profiles when exploiting a population with slow growth dynamics. The average profits for Player 1 (strategies in rows) given here are obtained numerically from a large number of iterations of the game, and are normalised to the value (0.0356) of the Nash equilibrium (ALLD vs ALLD) solution. Parameter values are $r = 0.5$, $K = 1$, $B = 1$, and $c = 0.2$.

will provide a link by which we can compare successful strategies from these deterministic game theoretic models with any results obtained when there is noisy population growth. In fisheries, the majority of noise occurs as variability in the reproductive process (Hjort, 1914; Swearer et al., 1999; Fogarty et al., 1991), and so we want to use a population model for which we will be easily able to add a random variable to the change in population. Furthermore, in order to be able to consider distinct generations of the population we want to use a population model for which spawning occurs at discrete time intervals, which represent the seasonal spawning of the population. As in the equilibrium game model, we look at the effects of a single fisher exploiting the population, under the same assumptions. Though reproduction occurs in discrete generations, the harvesting of the population will occur continuously throughout the fishing period, such that $u_n(t)$ is the population density at time t $(0 \leq t \leq T)$ during the nth period. When a single fisher fishes with constant effort F over the nth fishing period of length T, the fish population will experience exponential decay during this time, that is, $u_n(T) = u_n(0)e^{-FT}$.

We therefore consider the Beverton-Holt map $u_{n+1}(0) = f(u_n(T))$, which gives the new population biomass after spawning (at the beginning of a new time period) as a deterministic function of the biomass at the conclusion of the previous time period:

$$u_{n+1} = \frac{K u_n e^{(-FT)}}{u_n e^{(-FT)} + \left(K - u_n e^{(-FT)}\right) e^{(-rT)}} \tag{3.33}$$

which can be understood as a discrete-time equivalent of the previously considered logistic growth model. This equation describes the new population due to stock recruitment, with a time T between generations spawning, in which resource depletion due to fishing occurs.

3.6.1 Monopolist solution

Again, we begin by examining the profit maximising fishing behaviour for a single fisher, with exclusive access to the fishery. For this system to be in an equilibrium state we must

have:

$$u_{n+1} = u_n \tag{3.34}$$

so that we find equilibrium states when $u^* = 0$ or:

$$u^* = K\frac{1 - e^{((F-r)T)}}{1 - e^{(-rT)}} \tag{3.35}$$

Within a single time period, the fish biomass decays exponentially with fishing effort according to:

$$u(t) = u^* e^{-FT} = K\frac{1 - e^{(F-r)T}}{1 - e^{(-rT)}} e^{-FT} \tag{3.36}$$

The fisher acts according to the assumptions set out in Section 2.1, and so as in that section we can obtain a monopolist effort F^M, which is the action that will result the best output for that player by optimising their payoff $P^*(F^M)$:

$$\begin{aligned} P^* &= B\int_0^T Fu(t)\,dt - cFT \\ &= BF\int_0^T Ke^{-Ft}\frac{1 - e^{(F-r)T}}{1 - e^{-rT}}dt - cFT \\ &= \frac{BK}{1 - e^{-rT}}\left(1 - e^{(F-r)T}\right)\left(1 - e^{-FT}\right) - cFT \end{aligned} \tag{3.37}$$

to find the individual profit maximising value of fishing effort by numerically solving:

$$\frac{dP}{dF} = BKT\frac{\left(e^{-FT} - e^{(F-r)T}\right)}{1 - \exp(-rT)} - cT = 0 \tag{3.38}$$

3.6.2 One-stage two-player game

As in the earlier sections of this chapter, we will now develop the game corresponding to this separated time-scales model to allow the decision making behaviour of two players. We can begin by determining an equitable "Pareto solution" level of effort $F^C = F^M/2$ for each of the two (symmetrically behaving) players, which corresponds to the decision to cooperate in

a standard prisoner's dilemma. Similarly we can obtain the Nash effort F^D by noting that when two fishers are exploiting the population, we are exponentially depleting the population by the total effort $F_i + F_j$, so that:

$$u_{n+1} = \frac{K u_n \exp(-(F_i + F_j)T)}{u_n \exp(-(F_i + F_j)T) + (K - u_n \exp(-(F_i + F_j)T)) \exp(-rT)} \tag{3.39}$$

and following an identical process to that done in the monopolist version of this model. The profit obtained by Player i, for a given level of fishing effort F_j by the other player is:

$$\begin{aligned} P_i(F_i, F_j) &= B \int_0^T F_i u(t)\, dt - cF_iT \\ &= BK \frac{F_i}{F_i + F_j} \left(\frac{1 - e^{(F_i+F_j-r)T}}{1 - e^{-rT}} \right) \left(1 - e^{-(F_i+F_j)T}\right) - cF_iT \end{aligned} \tag{3.40}$$

With the population growing according to a Beverton-Holt map, we obtain the "Nash solution" effort F^D by solving numerically Equation (3.40).

3.6.3 Iterated two-player game

Finally, we again allow our single-shot game with both players to be repeated over an infinite time horizon, with each player choosing a strategy profile determining the actions they will take at stages of the iterated game, based on the state of the system. Without loss of generality, we let the carrying capacity $K = 1$, as well as the length of the fishing period $T = 1$. Letting $B = 1$ and $c = 0.1$, we investigate how the possible effort actions change as we alter the growth rate parameter r. Figure 3.4 (a) gives the equilibrium solutions to Equations (3.38) and (3.40), and shows that as expected, the choice to defect means exploiting the fish population more intensively than the choice to cooperate (as seen in each of the previous models). These special levels of effort which the fishers harvest at are used to determine the equilibrium population density under each possible outcome, or combination of effort choices by the players. Figure 3.4 (b) shows how the equilibrium density changes as the growth rate r increases. We note that for each possible value of r, the solution where

both players cooperate has the largest equilibrium density out of all possible outcomes, as expected when the total magnitude of exploitation of the fishery is the least. Figure 3.4 (c) shows that that the payoffs (profits to the individual players) for each of the possible outcomes are such that that the game constructed under the assumptions of this model satisfies the necessary condition to be a prisoner's Dilemma. Furthermore, Figure 3.4 (d) shows that the relative magnitudes of these payoffs are such that for all values of r the single-shot when repeated will be an IPD, with all corresponding properties of that game applying to it. As these results hold true as r is varied, we expect that the population growth rate parameter does not qualitatively affect the structure of the resulting game, and so perfect information strategies should have payoffs as expected in line with Section 3.4.4.

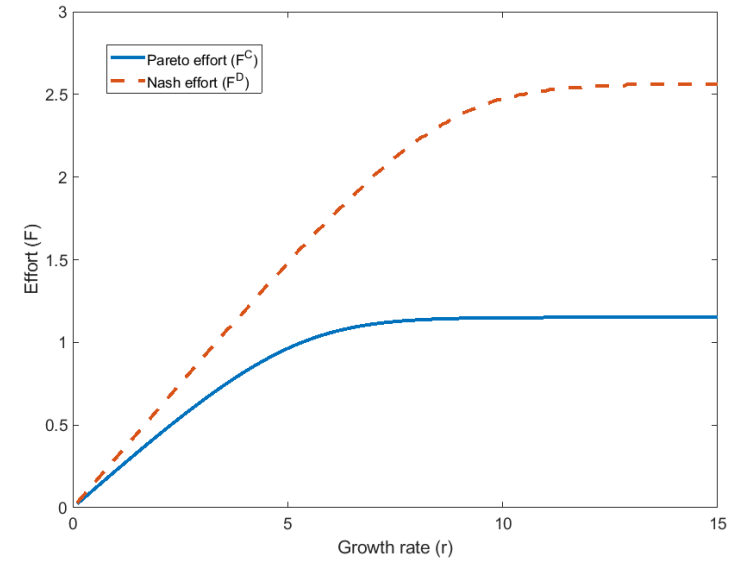
As with both the fast and slow dynamics continuous time models, we can numerically calculate the average profits obtained by a player for each of our combinations of imperfect information strategies. As before, this is done using numerical simulations, to find the average profit obtained by players using each combination of strategies, over a large number of fishing periods. In this model, Euler's method used to solve differential equations, and the MATLAB command "fzero" was used to find the levels of effort maximising profit in Equations (3.38) and (3.40). In Figure 3.5, a Beverton-Holt growth model with fast dynamics (with $r = 20$ as in 3.2) is used. We see that in this case, the normalisation factor for profits is much lower than in the logistic growth model with identical parameter values. However, the relative structure of the performances of each strategy, matching typical IPD results, remains. We see that memory-one strategies have a similar performance against each other as in Section 3.4.4. In particular, we see again in this case that TFT performs well against most other strategies; the only one that dominates it is Sly, which is designed to take advantage of "nice" behaviour. We note that Sly performs worse against itself than TFT performs against itself. We also observe that in this case, unlike the slow dynamics model, TFT cannot be invaded by $ALLD$. In general, with this high growth rate result we expect populations to reach equilibrium in each iteration, and so expect the performance

of the imperfect information strategies in simulations to correspond to the results from a standard IPD.

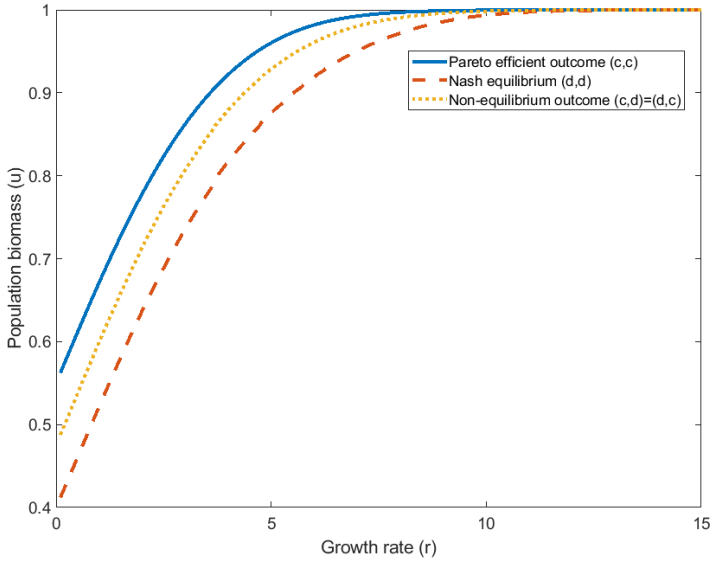
We can also compare the Beverton-Holt growth model with the logistic model when there are slow dynamics. Figure 3.6 shows the numerically determined average profits for this. With slow dynamics, the profit normalisation factor for the Beverton-Holt model is similar (but slightly lower) than that of the logistic growth case. In this model, we again see that $ALLD$ can beat TFT. However, unlike in Section3.5, TFT performs better against itself than $ALLD$ does against TFT.

3.7 Discussion

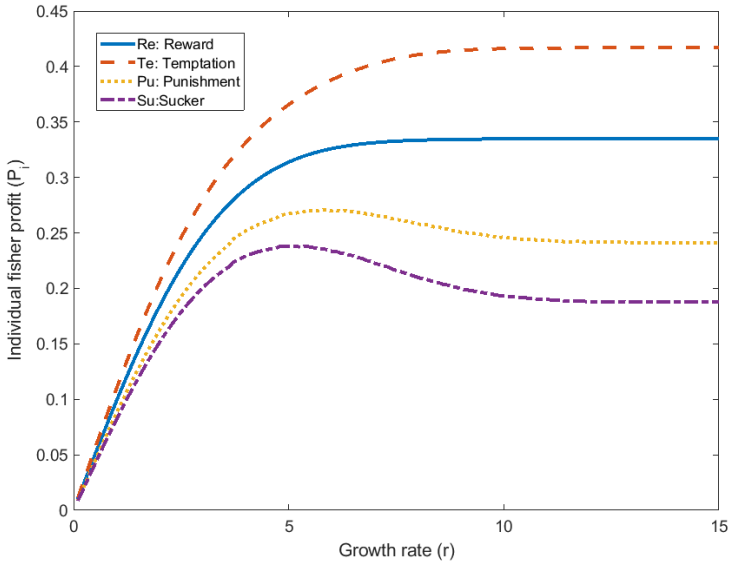
In the construction of our decision making model, fishers are solely motivated by their own payoffs. In particular, they wish to maximise their profit from the fishery, whether in a single fishing period in one stage-game, or over a period of time in the iterated game. If we instead take a more holistic view of the ecosystem, then we can consider how different types of fishing behaviour and strategy profiles affect not only the economic output of the fishery, but also the ecological impacts of these actions. In this model, the ecological health of the population is measured by considering the population biomass under exploitation. Specifically, we can compare the resultant population biomass for a specific harvesting behaviour to the unexploited biomass, to see the extent of the population depletion that occurs due to fishing. We note that when a prisoner's dilemma scenario emerges from our model, the Pareto optimal solution is associated with a lower aggregate fishing effort than the Nash equilibrium. In a Gordon-Schaefer population model, higher rates of exploitation result in lower abundance. As such, the Pareto optimal solution has a higher equilibrium population biomass than the Nash equilibrium. This means the Pareto equilibrium is not just socially optimal in the game theoretic sense of maximising the total payoffs obtained by the fishers, but also in terms of minimising the ecological impact of population depletion.



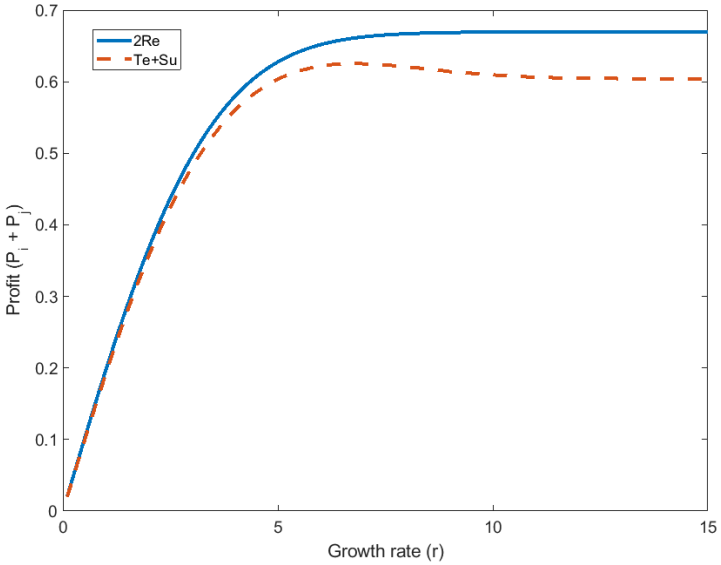
(a)



(b)



(c)



(d)

Figure 3.4

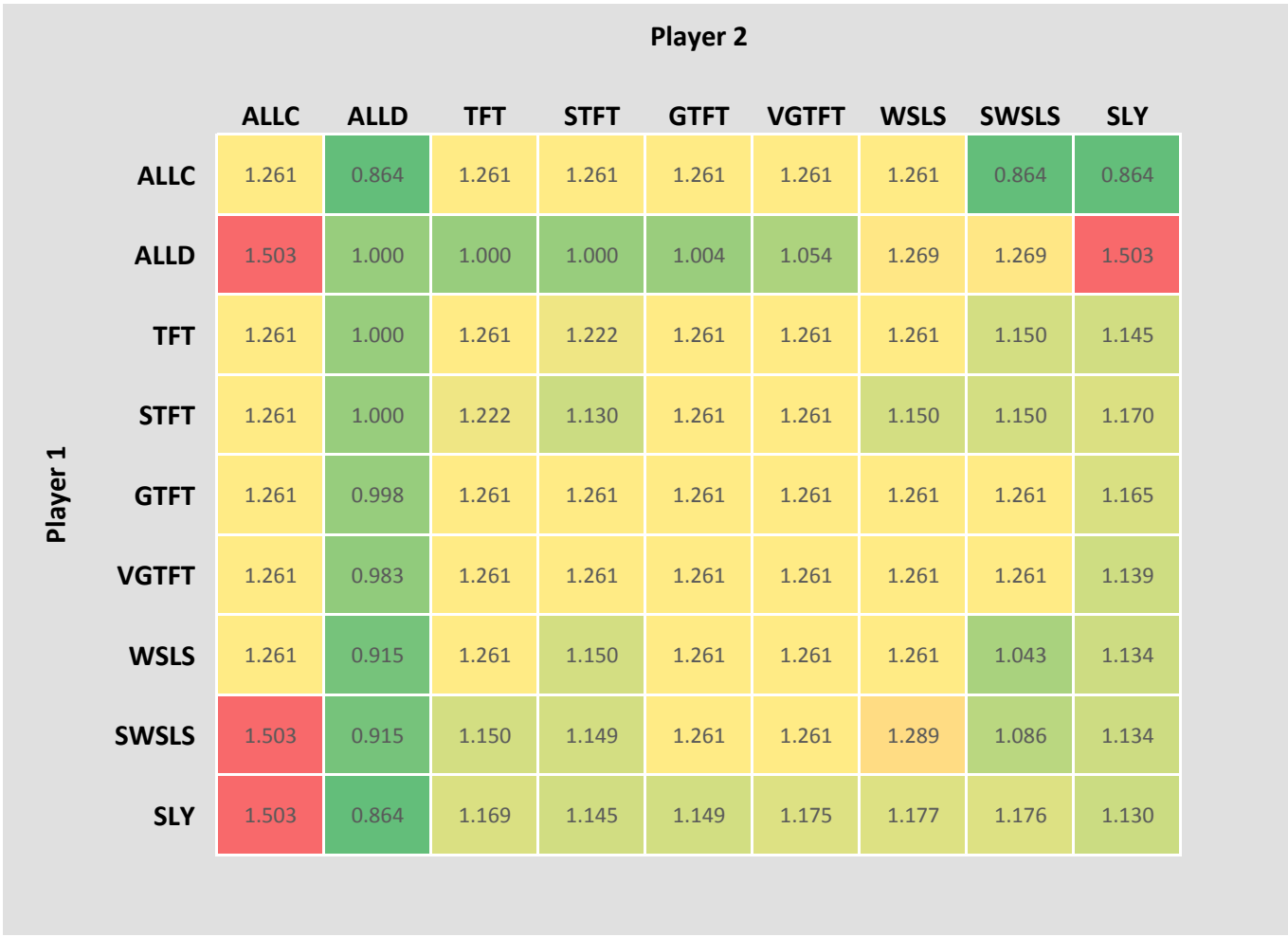


Figure 3.5: A strategy payoff heatmap using the Beverton-Holt population model with fast growth dynamics. Parameter values are $r = 20$, $K = 1$, $B = 1$, and $C = 0.2$. Payoff values are normalised by a factor of 0.1896, the value of the Nash equilibrium payoff. We note that the payoffs for each strategy profile are significantly lower than the corresponding logistic growth example with identical parameter values.

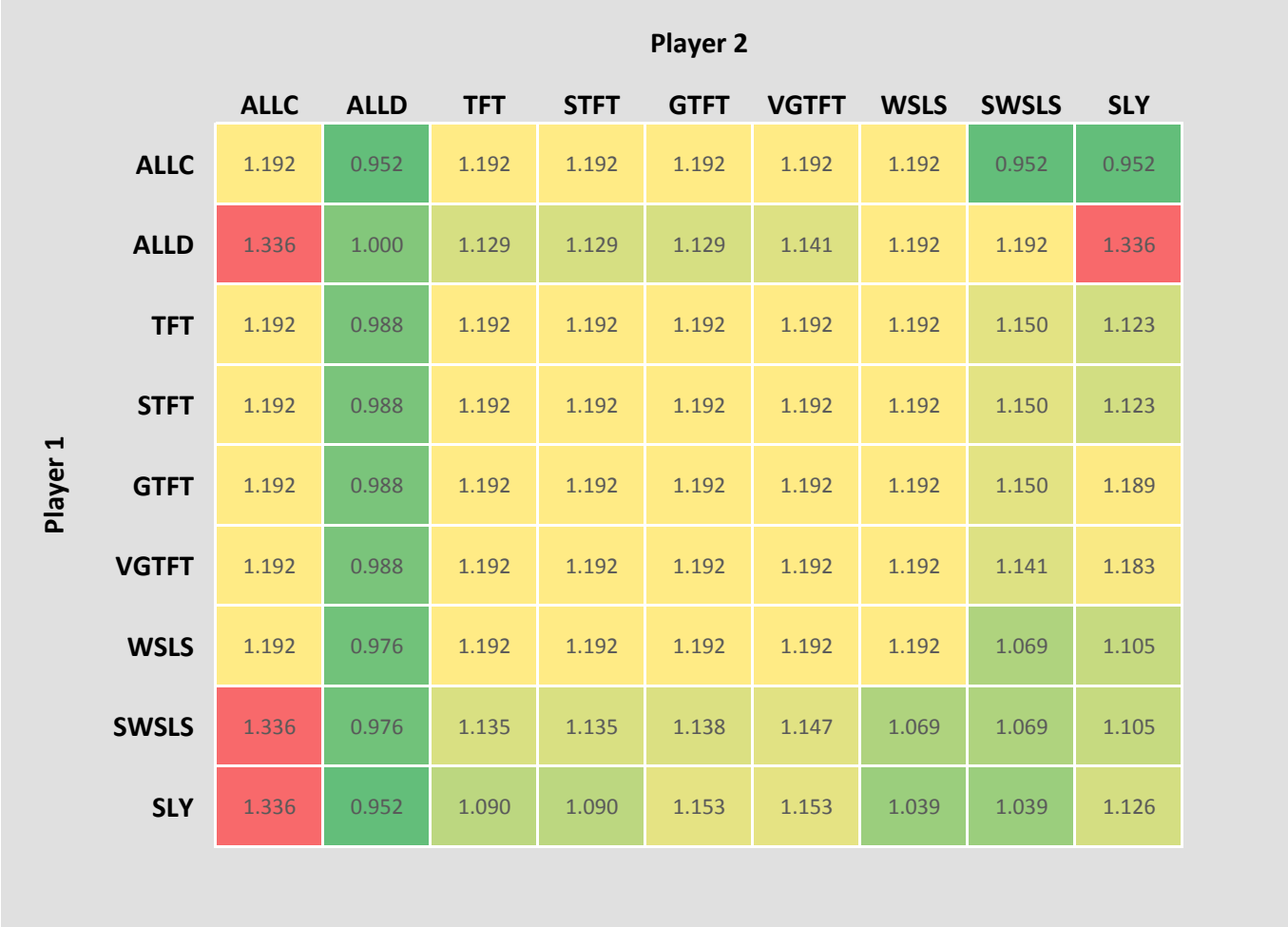


Figure 3.6: A strategy comparison heatmap for a population with Beverton-Holt growth with slow dynamics. Parameter values are $r = 0.5$, $K = 1$, $B = 1$, and $C = 0.2$. The normalisation factor here is 0.0333, the Nash equilibrium payoff. Payoff values are in each case slightly lower than in the corresponding logistic growth model.

In our fast resource dynamics model, we explicitly see in Equation (3.26) that in a two-player game, an increase in effort by one player will result in a reduction in the profit gained by the other player (when that player is active in exploiting the fish population). This is what is known as "the tragedy of the commons" (Hardin, 1968), where the costs resulting from increased exploitation are shared by all who have access to the resource, but the benefits are only received by the fisher who increased the intensity of their exploitation. While this concept is typically applied to open-access resources, which is not strictly the case in our two player model, it was originally formulated in terms of "common" grazing land which was shared by a small number of users. Furthermore if our model was extended to an open-access fishery with n fishers (which is simple to do in the fast resource dynamics case), this tragedy of the commons would still occur. There, the increased economic costs of over-exploitation would be shared by all fishers, with the increased benefits only obtained by one. However, the concept of the actions (due to increased exploitation) of one fisher imposing an increased economic cost on another fisher still applies. In our model, this higher economic cost occurs due to a reduction in the population biomass available to be caught by the other player. There has been some criticism of the tragedy of the commons as a concept, with some contending that it does not accurately describe real life situations (Feeny et al., 1996). Indeed, the fact that common land was successfully used for centuries without critical overexploitation occurring has led to the suggestion that common property resources result in "triumph", not tragedy (Cox, 1985; Coop and Brunckhorst, 1999). In many real life fishing situations the tragedy of the commons has been observed to not occur, as decision makers collectively act in ways to avoid negative outcomes, and protect their resource (Berkes, 1985).

In our model, we are trying to understand the decision making behaviour of fishers interacting without regulation or intervention from any overriding governing body. It is therefore not appropriate to treat our results as situations that can be dealt with by adding additional factors modelling such regulation. However, in our repeated games, we see that coopera-

tive solutions can arise naturally between players, with cooperative strategies such as TFT performing well in terms of payoffs. In our cooperative outcomes, we saw that individual players acting for their own best interest in the short term (by acting to maximise profits) resulted in outcomes that were in the collective best interest for both players (in terms of both total profit at the Pareto solution, and minimised population depletion).

Pareto optimality occurs when it is impossible to allocate resources, through the choice of the fishers' efforts, in such a way that makes any player better off (in terms of increased payoff) without making the other player worse off. However, a system being Pareto efficient does not specify that any degree of equality or equitability has occurred. In our model, we defined our "cooperative" or Pareto effort to be that of half the monopolist effort. This is an example of a Pareto optimal strategy, and one in which both players will receive equal payoffs (due to their equal effort and costs). However, any situation in which $F_1 + F_2 = F_M$ holds true will in fact be a Pareto efficient solution, as in any such case an increase in effort by either player will result in a reduction in the payoff to the other player, meaning Pareto optimality would hold. For simplicity, and noting the symmetries of the players' situations, we made an assumption restricting effort choices to either the Nash equilibrium effort F_D or our single Pareto effort F_C, in order to investigate the type of two-player, two-action game that would arise (and in doing so determined this game was equivalent to a prisoner's dilemma). However, any of the other Pareto equilibria would still be valid solutions to the game. We do however note that, as seen in Equations (3.22) and (3.26), the rate of change of a player's profit as either player changes their actions depends on the relative size of those efforts. The potential increase in payoff to the player with the lowest payoff that could be obtained by choosing to defect and move away from the Pareto equilibrium will be minimised when $F_1 = F_2 = F_C$, relative to any of the unevenly distributed Pareto solutions. The player with the greater share of the profit (that is the higher effort) will instead have a smaller incentive to fish at the Nash effort, as the marginal increase in payoff due to defection will be lower for the fisher with the higher profit at an uneven Pareto solution.

The separation of time-scales in which reproduction and fishing occur is a reasonable assumption. Spawning periods often occur on an annual time-scale (Booth and Beretta, 1994; Allen and Barker, 1990; Campana, 1996), whereas fishing can occur either seasonally or continuously throughout the year. In either case, the spawning period can occur over a much shorter period of time than the length for which fishing occurs (Conover and Present, 1990). By discretizing the reproductive dynamics, the Beverton-Holt growth model is therefore suitable for this purpose. As it is the discrete time equivalent of a logistic model, it also captures similar behaviour to the Gordon-Schaefer model in terms of rates of growth relative to carrying capacity (Berezansky and Braverman, 2004). We also assumed that the rate of exploitation by a fisher was fixed at a constant value for the entire length of a single fishing period, with the fishers having the ability to change it between fishing periods. As the length of each fishing period is essentially an arbitrary one (which we generally allowed to be one year), this assumption is also reasonable, as we could change the length of the period in the implementation of the model in order to see how the ability to change behaviour at a higher frequency effects the behaviour of fishers and their payoffs. This will have a predominant impact on the slow resource dynamic case, as the underlying assumption of that model is that fishing and reproduction occur on similar time-scales.

There is significant scope to extend this modelling process further. As a starting point, there are an unlimited number of mixed strategies that can be used in a two-player game, though not all are either interesting or intuitive as a decision making process (Tadelis, 2013). We have mainly considered memory-one strategies, though strategies in which players have memories of any length could be developed and used. However, there is likely to be some computational trade-off between any potential gains in terms of payoff values, and the increased complexity of some strategies. Furthermore, while there may be an increase in obtained payoffs when playing against the current strategies, this does not preclude other new strategies or combinations of strategies from outperforming them (Boyd and Lorberbaum, 1987).

Particular, promising strategies could be chosen, or a method such as a genetic algorithm could be used to generate possible strategies that could be compared in simulations to our existing strategies (Axelrod, 1987).

Chapter 4

Applying Markov decision processes to the monopolistic exploitation of a fishery with stochastic stock recruitment

4.1 Introduction

4.1.1 Overview of chapter

Fish populations are noisy (Hjort, 1914; Doherty and Williams, 1988). In this chapter, we will describe optimal fishing behaviour in such noisy conditions, taking into account how both current decision making behaviour and biological variability will impact future profits. To do so, we use Markov models; fishing behaviour is considered "memoryless", depending only on the present state of the fish population. The first part of the chapter details the optimal harvesting behaviour for a monopolist. We use Markov decision processes (Puterman, 2014) to model this monopolist fisher's decision maker behaviour, obtaining an "optimal policy", mapping from the current population state to the best choice of action, for specific

model conditions. Next, this concept is extended to a two-player game, where each fisher is independently attempting to maximise their own long-term profit. This Markov game (Littman, 1994) is solved to find a policy for each player that is a "best-response" to their beliefs of how the other will harvest the fish population.

The noise in fish stock levels can occur because of environmental reasons such as: seasons, climate (Lehodey et al., 2006) and temperature variations (O'Brien et al., 2000); inter-specific interactions (Pritt et al., 2014); variations in available nutrients (Swearer et al., 1999); spatial heterogeneity (Wiens, 1976; Levin, 1976); simple differences in individual characteristics (Lomnicki, 1978; Tyler and Rose, 1994); and fecundity of members of the population (Houde and Hoyt, 1987; Pepin and Myers, 1991). Noise can be expressed both continuously over time through growth and death processes, and discretely in the reproductive process. It is this variability in stock recruitment that has the predominant impact on the stochasticity of population levels (Hjort, 1914; Swearer et al., 1999; Fogarty et al., 1991), and which will be investigated in this chapter.

Generational spawning is common in fish species, with this often occurring annually (Booth and Beretta, 1994; Allen and Barker, 1990; Campana, 1996) over short spawning seasons (Conover and Present, 1990). Stock recruitment relationships are used to describe the growth of a fish population from generation to generation, with concern paid to the ways in which the current level of a fish population will be responsible for future population levels (Beverton and Holt, 2012; Ricker, 1954).

4.1.2 Population modelling

One commonly used deterministic model of the stock recruitment relationship is the Beverton-Holt growth model (Beverton and Holt, 2012), which can be considered to be the discrete time equivalent of a logistic (Pearl-Verlhust) growth process (Berezansky and Braverman, 2004). The Beverton-Holt map gives the expected density u_{n+1} of the population in genera-

tion $n+1$ as a function of the expected density u_n in the previous generation. The population is modelled to have a fixed carrying capacity K, a constant growth rate r, and with time T between generations. The new expected population density is then given by

$$u_{n+1} = \frac{Ku_n}{u_n + (K - u_n)\exp(-rT)}. \tag{4.1}$$

Other alternative methods of modelling fish populations also exist. For instance, the Ricker model (Ricker, 1954), in which

$$u_{n+1} = u_n \exp\left(r\left(1 - \frac{u_n}{K}\right)\right) \tag{4.2}$$

is also a discrete time model of successive generations. It differs from the Beverton-Holt model in that at high stock sizes, recruitment will begin to decrease, in order to capture the effects of factors such as cannibalism. Both the Beverton-Holt model and the Ricker model contain both density-dependent and density-independent terms. Continuous time models such as the logistic growth model can also be used. For such a model, any noise could be added as a stochastic process to create a stochastic differential equation (or could be added in yearly increments). Continuous models could be used to represent sources of noise that occur consistently throughout the individual growth processes occurring between generational spawning.

4.1.3 Markov decision processes

When mathematically modelling the decision making behaviour of individuals in situations where both the decisions themselves and randomness can affect outcomes, Markov decision processes (MDPs) are an effective tool to use (Puterman, 2014). MDPs are discrete time stochastic control processes consisting of

- A state space S
- A set of actions A which the decision maker can choose from

- For each pair of states s and s', and action a, the transition probability $Q(a, s, s')$ which is the probability that the choice of action a in state s will lead to state s'.
- $R(a, s, s')$ is the *immediate* reward (or one-stage payoff) for the decision maker when a transition from state s to state s' occurs, due to the choice of action a.
- δ is the discount factor of future rewards compared to current rewards, equivalently the fraction of the profit this year that an equal profit next year is valued.

MDPs exhibit the Markov property, in that the future effects of any decision at any time only depend on the current state of the system, without any dependence at all on the history of the system (Puterman, 2014). As such, MDPs are an extension of a Markov process to allow the influence of actions made by a decision maker on the outcomes.

Markov decision processes are solved by finding an optimal policy $\Pi : S \to A$, which is a mapping from the set of states S to the set of actions A, and which specifies the action a decision maker should take in each state, at each time (Puterman, 2014). In a single time period, a decision rule specifies which action should be taken for each state. The collection of all decision rules over time forms a policy (Puterman, 2014). That is, a policy Π is comprised of a series of decision rules, one for each decision epoch in the time horizon. Here we will use subscripts to denote the specific decision rule that is followed at a given time. Therefore $\Pi_n(s)$ will denote the decision rule for the policy Π in the decision epoch n. A policy is optimal in the sense that it maximises the expected total discounted reward that the decision maker will receive over the entire time horizon. That is, in solving the MDP we want to find a policy Π such that setting $a_n = \Pi(s_n)$, we can maximise the expected value of:

$$\sum_{n=0}^{\infty} \delta^n R(a_n, s_n, s'_{n+1}) \tag{4.3}$$

A policy Π is called "stationary" if it does not change with time, so that the choice of action depends only on the state, and not on the decision epoch of the MDP itself (Puterman, 2014). A necessary condition for the optimality of a policy (Dixit et al., 1990) is that it

satisfies the Bellman equation (Bellman et al., 1954). This equation recursively defines a value function V for the decision problem at a specific time, in terms of the current one-stage payoff obtained due to the choice of some action, and the expected value of being in the resulting state, based on that action (Bellman, 1966). The value $V_n(s_n)$ is then the expected discounted reward from time n until the end of the (possibly infinite) time horizon, given that the state is s_n at time n. The Bellman equation is

$$V_n(s_n) = \max_{a \in A} \{R(a, s_n, s'_{n+1}) + \delta \sum_{s' \in S} Q(a, s_n, s'_{n+1}) V_{n+1}(s'_{n+1})\} \tag{4.4}$$

A Bellman equation reduces a decision problem into smaller subproblems, according to Bellman's "Principle of optimality", that "An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" (Bellman, 1966). This is analogous to the concept of subgame perfect equilibria in dynamic game theory. Bellman equations are particularly suited to solving stochastic optimal control problems, such as Markov decision processes. Intuitively, the Bellman equations looks to maximise the sum of the current and future reward available from a given state, by choosing the best possible action.

4.1.4 Solution methods for Markov decision processes

Some common solution methods for an infinite time horizon problem are value iteration and policy iteration, each of which act by performing updates on the Bellman equation to compute optimal values (Puterman, 2014). Value iteration aims to find an optimal value function, and then performs one policy extraction, to find the policy associated with the optimum value. With this technique, the policy corresponding to the value function only needs to be found once, because if the value function is optimal, then the policy will also be optimal. A general value iteration algorithm given by Puterman (2014) is:

1. Initialize a value function V^0, specify a tolerance $\epsilon > 0$, and set the iteration $n = 0$.

2. For each $s \in S$, compute:

$$v^{n+1}(s) = \max_{a \in A} \{R(a, s, s') + \delta \sum_{s' \in S} Q(a, s, s') V^n(s')\} \tag{4.5}$$

3. If

$$||V^{n+1} - V^n|| < \epsilon \left(1 - \delta\right) / 2\delta \tag{4.6}$$

proceed to the next step. If not, increase n by 1 and return to step 2.

4. For each $s \in S$, let:

$$\Pi_\epsilon(s) \in \arg\max_{a \in A} \{R(a, s, s') + \sum_{s' \in S} Q(a, s, s') V^{n+1}(s')\} \tag{4.7}$$

and stop.

In the above algorithm, steps 2 and 3 attempt to find the optimal value function for the problem. Step 4 is the policy extraction, giving the policy corresponding to this value.

Policy iteration involves an evaluation of the current policy, and then policy improvement. This is then iterated until convergence occurs. Puterman (2014) gives a policy iteration algorithm as:

1. Initialise the iteration $n = 0$, and select an arbitrary decision rule Π^0 for this iteration

2. Solve:

$$\left(I - \delta \sum_{s' \in S} Q(a, s, s')\right) V = R(a, s, s') \tag{4.8}$$

for V^n.

3. Choose Π_{n+1} that satisfies:

$$\Pi^{n+1} \in \arg\max_{a \in A} \{R_\Pi + \delta \sum_{s' \in S} Q(a, s, s') V^n\} \tag{4.9}$$

selecting $\Pi_{n+1} = \Pi_n$ if possible.

4. If $\Pi_{n+1} = \Pi_n$, stop and let the stationary policy $\Pi_* = \Pi_n$. Otherwise, increment n by 1 and go to step 2.

Here, step 2 is the policy evaluation step, and step 3 is policy improvement. In terms of use, policy iteration often requires fewer iterations than value iteration, particularly if the transition probabilities for the problem are structured as a sparse matrix. However, policy iteration is more computationally expensive, as it requires solving a linear system rather than simply applying a linear operator.

A value iteration dynamic programming technique for a finite time horizon problem is to use backwards induction on our stochastic dynamic programming equation. This is a recursive method, in which each time period of the Bellman equation is solved in reverse order. The final period is a straightforward reward maximisation problem, as there is no future value in the system at this point (although depending on the formulation of the problem there may be some "salvage value" of remaining stock (Puterman, 2014)). Each preceding time period is solved iteratively, taking into account the (known) future value of being in every state, so that the effectiveness of each possible choice of action can be determined. At each time period, the best choice of action for each state is determined, and stored as an optimal policy.

An infinite time horizon is the limiting value of the finite time horizon scenario, as the number of time periods goes to infinity. If the finite time horizon considered is sufficiently long that the optimal policy converges to a stationary (time-invariant) one before the final iteration, then that optimal policy can be considered equivalent to a stationary optimal policy under an infinite time horizon. If the discount factor δ is less than one, the total expected reward value will converge to a finite solution if the time horizon is large enough (Sutton and Barto, 2018). Furthermore a stationary optimal policy is guaranteed to exist if the state space and action space are both finite, or equivalently if all single-stage rewards are bounded (Puterman, 2014). Such a stationary optimal policy is not guaranteed to be unique, but at

least one will be deterministic (Howard, 1960).

There are some downsides to the use of MDPs. The so-called "curse of dimensionality" is that as the state space and action space become large, MDPs get very computationally difficult to solve (Bellman et al., 1954). There are also large memory requirements, as for every pair of an action and state, we require both a transition probability matrix and reward function. Solutions also rely on the assumption that transition probabilities are constant over time, which may not necessarily be realistic. There are approximate solution techniques which may help overcome some of these problems. These techniques do not find an exact solution, but may be faster. Examples of these include: state aggregation, in which several qualitatively similar members or regions of the state space are grouped together, reducing computation time by reducing the repetition of similar outcomes; action elimination, where redundant or unlikely actions are excluded from the action space; and approximate dynamic programming (Powell, 2007).

4.2 Discrete toy model

How much fish should a monopolist catch? If there is no concern for the future, they will simply want to maximise profits in the current year. But if the individual is cognisant of the fact that their own actions (the effort with which they attempt to exploit the fish population) will deplete the stock, and therefore, impact possible profits in the future, they may act differently. If they have the ability to fish into the future, then they will want to maximise the value they can obtain from the fishery over that entire period. This may be a long time, or indeed there may be an infinite time horizon. So how do we deal with these potential future profits? MDPs provide a useful framework for dealing with such a problem. Our main objective in this chapter is to develop a model in which we are able to use a MDP in order to obtain an optimal strategy for harvesting a stochastic renewable fish population. To get to this point, we will first use a simpler toy model to highlight necessary features of

this application.

4.2.1 Deterministic decision problem

To illustrate the MDP framework, we will begin by considering a toy model of a fish population, which will allow us to depict the key features of our solution method without needing extensive numerical calculations. Our "toy" fish population has its density measured by stratifying population levels into three categories: "High", "Medium", and "Low". Fishing depletes populations, and so if it occurs we expect biomass to decrease. If the density is high and fishing occurs, it will decrease to a medium level; similarly fishing will cause a medium density population to become low density. The fish population will also naturally grow, and so in the absence of exploitation, population levels will replenish. If the population is at low density and no fishing occurs, it will reach medium density, and if it is at medium density, a lack of exploitation will result in it reaching high density. The fish population cannot reach a density above the "high" level, even in the absence of harvesting. This level represents the natural carrying capacity of the population. This information is captured in Figure 4.1, a schematic diagram showing each possible population state, and the allowable transitions between them due to particular actions.

An individual who has the opportunity to exploit the population must therefore make a choice whether they fish (taking action "F") or do not fish (taking action "DF"). The information which will influence their decision comes solely from the state of the system, which is its population density at that time, without any influence from the previous states of the population. The system therefore exhibits the Markov property of memorylessness.

The allowable transitions between states under each action can also represented as a table of transition probabilities from state i at time n to state j at time $n+1$. With deterministic growth and fishing behaviour, allowable transitions will occur with probability 1, and state changes that cannot occur have probability 0. Table 4.1 gives the transition probabilities

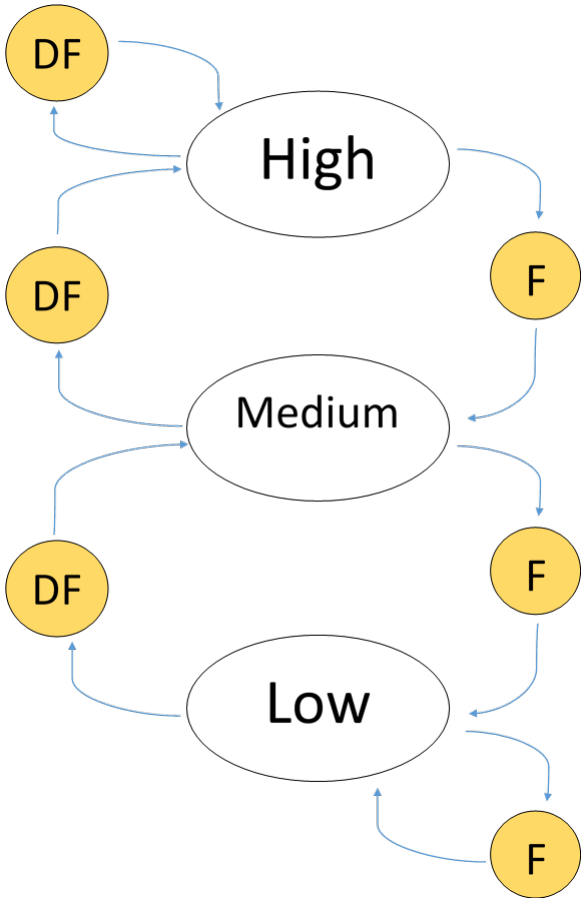


Figure 4.1: State space diagram for all possible transitions in the deterministic toy model of the fishery. White ovals show the possible population density states: "High", "Medium", and "Low". Yellow circles show the possible effort decisions at each state: "Fish" (F) and "Don't Fish" (DF).

F		$n+1$		
		High	Medium	Low
	High	0	1	0
n	Medium	0	0	1
	Low	0	0	1

DF		$n+1$		
		High	Medium	Low
	High	1	0	0
n	Medium	1	0	0
	Low	0	1	0

Table 4.1: State transition matrices for the deterministic optimisation problem, for both of the actions available to the decision maker at time n.

for each combination of state and action. We can also assign numerical values to the profits obtained by the decision maker when state transitions occur. If they decide not to fish at time n, then they obtain no yield from the fishery and thus obtain no profit in the current time period. We also assume that they will obtain a higher profit from fishing when the population is at high density than at medium density, and similarly a higher profit when fishing at medium density than at low density. Table 4.2 gives an example of possible profit levels from each state transition when fishing under these conditions.

We then have the following at each decision epoch:

- A state space $S := \{\text{High},\ \text{Medium},\ \text{Low}\}$
- An action space $A := \{\text{Fish, Do not Fish}\}$
- A set of transition probabilities $Q(a, s, s')$ from state $s \in S$ at time n to state $s' \in S$ at time $n+1$ (Table 4.1)
- The immediate rewards gained from taking an action $a \in A$ and transition from state s to state s' (Table 4.2).
- No discounting ($\delta = 1$)

Formally, we wish to solve the deterministic value problem

$$V_n(s) = \max_{a \in \{F, DF\}} \{R(s, a) + \sum_{s' \in S} Q(a, s, s') V_{n+1}(s')\} \tag{4.10}$$

F		$n+1$		
		High	Medium	Low
	High	0	7	0
n	Medium	0	0	5
	Low	0	0	1

Table 4.2: One-stage reward matrix when the decision maker chooses to fish at time n. Note that the one-step reward is 0 for any state transition if the decision maker chooses not to fish.

$n=4$	One-step reward $R(s,a)$		Optimum	
State	F	DF	$V_n(s)$	$\Pi_n(s)$
High	7	0	7	F
Medium	5	0	5	F
Low	1	0	1	F

Table 4.3: Outcomes for each choice of action during the final time period $n = 4$

Doing so will give us an optimal policy $\Pi_n : S \to A$.

As an illustrative example of the backwards induction process used to determine an optimal policy, we work through the decisions that should be made in a case where the fisher obtains the profits given in Table 4.2, and will fish for a finite period of four years. Here we will use the notation $\underline{s}$ = (High, Medium, Low), such that in year n we can describe a full policy as $\Pi_n(\underline{s}) = (a_H, a_M, a_L)$, where for example $a_H = \Pi_n(H) \in \{F, DF\}$ is the optimal action the fisher should take if the population level is high. We will also use $\underline{V}_n(\underline{s})$ =(V_n(High),V_n(Medium),V_n(Low)) to denote a vector of the values of being in each

$n=3$	One-step reward $R(s,a)$		Optimum	
State	F	DF	$V_n(s)$	$\Pi_n(s)$
High	12	7	12	F
Medium	6	7	7	DF
Low	2	5	5	DF

Table 4.4: Outcomes of each choice of action during the time period $n = 3$

$n = 2$	One-step reward $R(s, a)$		Optimum	
State	F	DF	$V_n(s)$	$\Pi_n(s)$
High	14	12	14	F
Medium	10	12	12	DF
Low	3	7	7	DF

Table 4.5: Outcomes of each choice of action during the time period $n = 2$

$n = 1$	One-step reward $R(s, a)$		Optimum	
State	F	DF	$V_n(s)$	$\Pi_n(s)$
High	17	14	17	F
Medium	12	14	14	DF
Low	4	10	12	DF

Table 4.6: Outcomes of each choice of action during the time period $n = 1$

possible state at time n.

We begin by looking at the final time period, where we only need to worry about current (and not possible future) profits. In the last year ($n = 4$), fishing results in a profit of 7 when density is high, 5 when medium, and 1 when low. Choosing not to fish results in zero immediate reward for each density (as it does at any decision epoch). In each case we see that it is preferable to fish than to not fish in the final time period, regardless of the current state of the population, because preserving population levels by abstaining from fishing will not be economically beneficial without any further opportunities to exploit the fishery. Thus, as shown in Table 4.3, we have a policy $\Pi_4(\underline{s}) = (F, F, F)$, which gives the fisher a value $\underline{V_4}(\underline{s}) = (7, 5, 1)$. In the penultimate year ($n = 3$), the decision maker must take into account what the outcome will be in the next year because of their choice whether to fish or not. The options available to the fisher, and their possible payoffs, are summarised in Table 4.4. If the density is high at $n = 3$, choosing to fish will result in a one-time profit of 7 in that year. It will also cause the population to be depleted to medium density, from which a profit of 5 is available in the last year (as noted above) giving a total future value of 12 obtainable from the choice to fish. On the other hand, choosing not to fish at $n = 3$

will result in an initial profit of 0 in that year, with the population density recovering to a high density in the next year, from which a profit of 7 is available. Choosing not to fish at $n = 3$ and with high density will therefore enable a combined value of 7, and so it is clearly optimal for the decision maker to choose to fish at $n = 3$ when density is high.

If the population density is medium during this second to last year, then a different situation will occur. Choosing to fish will result in an immediate profit of 5, with the optimal future profit being the 1 gained by fishing again in the next year, for a total value of 6. Choosing not to fish will mean no profit during $n = 3$, with the population then being replenished to a high density for the final year. As seen in Table 4.4, the best decision then is to not fish, such that the fisher will obtain a future value of 7. Consequently at medium density, not fishing is the optimal policy. A similar situation occurs at $n = 3$ with low density. Choosing not to fish at this time allows the population level to increase in the next time period. The gain in value (a profit of 5) from allowing this to happen (and then fishing during the final year) is greater than the combined small profits (1 in each year for a total of 2) received from fishing in each of years 3 and 4. Therefore with low density the best policy is also to not fish at $n = 3$. Concluding, we see that the optimal policy in this year is $\Pi_3(\underline{(s)}) = (F, DF, DF)$, which means in each possible population state the agent has value $\underline{V}_3(\underline{s}) = (12, 7, 5)$).

We now repeat the process for the previous year ($n = 2$), as summarised in Table 4.5. In this year, the optimal policy is again $\Pi_2(\underline{(s)}) = (F, DF, DF)$, which in this case gives the fisher a value of $\underline{V}_2(\underline{s}) = (14, 12, 7)$). Looking at the final stage (that is, the first year of exploitation), which is detailed in Table 4.6, we see that the optimal policy is to fish when the population density is high, and not to fish when it is medium and low, $\Pi_1(\underline{(s)}) = (F, DF, DF)$ with optimum value $\underline{V}_1(\underline{s}) = (17, 14, 12)$). We note that this is the same policy at $n = 1$ as at $n = 2$ and $n = 3$, meaning that the policy has converged. We have therefore found a stationary optimal policy for the harvesting of this fishery, that will be followed in each year

of exploitation except the final one. Consequently, if the finite time frame was increased to longer than 4 years, we already know the optimal policy at all times, and can indeed extrapolate to a solution to an infinite time horizon problem. Specifically, in the infinite horizon regime the optimal policy is to only fish when the density is high.

4.2.2 Stochastic decision problem

Real fish populations exhibit noise in their population levels. In particular, the majority of this noise will occur in the stock recruitment relationship, with the number of fish spawning in a given year being partially, but not entirely, dependent on the spawning stock biomass. With a noisy population finding an optimal policy is a stochastic dynamic programming problem, meaning MDP techniques are able to be used to determine the optimal policy when an individual is harvesting the fish population. We extend the toy model from Section 4.2.1 to illustrate how the process is adapted to allow for this stochasticity. In order to incorporate noise in the dynamics, we will now allow all possible transitions between the states to occur with each action, with different specific transitions occurring with different probabilities. This stochasticity in the transition probabilities is used to represent the natural variability in the growth processes of the fish population, and in particular noise in the stock recruitment relationship. The probabilities of these transitions are arbitrary for the purpose of this illustrative model, with their values chosen such that more reasonably realistic transitions will be more likely to occur. For example, if the current state of the population is medium density and fishing occurs, then intuitively the most likely transition is to low density at the next iteration, as the impact of the population reduction due to exploitation is more significant than the natural growth of the remaining population. However, there is some chance that the size of the effects is similar, and so the population will remain at medium density at the next iteration. There is also some small chance that the fish population will have an unexpected bumper year, with the growth processes resulting in a large level of stock recruitment. Then the fish population could be at a high level in the next year, even with fishing occurring. In each of these cases, a consistent level of fishing effort

will have occurred, and so the fisher will obtain the same yield, and receive the same profit, as if the transition to a low density had occurred. Therefore the transition that actually occurs is random, but the immediate profit obtained by the decision maker is deterministic based on the action they choose and the existing state of the fishery. The new transition probabilities for either choice of action are given in Table 4.7.

Immediate profits for the decision to fish at each population density are still as given in Table 4.2. We note that while the immediate profit is only dependent on the current state and the action taken, these actions themselves can now lead (with known probability) to quite different outcomes. Consequently, we consider both the immediate profits and the expected future value from a decision instead of the deterministic future value when evaluating that decision, in order to determine the optimal policy which will maximise the objective function in a given iteration.

As before, we attempt to ascertain an optimal policy for harvesting this population over a period of four years. Though transitions are now probabilistic, in the last year of exploitation the decision maker is not concerned with future profits and so will again maximise immediate profit in the year $n = 4$. As such, the optimal decision for each density level in the stochastic model is the same as in the deterministic model, with it being preferable for the decision maker to fish at all density levels ($\Pi_4(\underline{s}) = (F, F, F)$), as seen in Table 4.8. The decision making information for the penultimate year $n = 3$ is captured in Table 4.9, where we see that as in the deterministic example the fisher will wish to maximise the sum of their immediate profit and the future value that will be available in the fishery. However, in this case the efficacy of an action is measured in its expected value, and so both the likelihood of each outcome as well as the actual payoffs resulting from these outcomes must be taken into account. For instance, when fishing at high density there is a 10% chance of remaining at high density in the last year, which is valued at 7. There is an 80% chance of the population being at medium density in the last year, which has a value of 5, and a 10% chance of the

population being depleted to low density, which has a value of 1. As such the expected future value of choosing to fish at high density is $0.1 \times 7 + 0.7 \times 5 + 0.1 \times 1 = 4.8$. If the decision maker chooses not to fish, it is certain that the population will remain at high density, and so the expected future value is 7. We conclude that at $n = 3$ and at high density the decision maker should fish.

Looking at a medium density population in the penultimate year, with the decision to fish there will be a 5% chance of the population reaching high value even in the event of fishing, due to an unexpectedly prolific spawning season. There is a 15% chance that the population stays at medium density, valued at 5 in the next year, and a 80% chance the population is at low density, valued at 1. The expected future value of choosing to fish at low density is $0.05 \times 7 + 0.15 \times 5 + 0.8 \times 1 = 1.9$. If fishing does not occur, there will be a 90% likelihood that the population density increases to a high level, and a 10% chance it remains at medium density. This decision will therefore have an expected future value of $0.9 \times 7 + 0.1 \times 5 = 6.8$, and we see that the optimal policy to maximise total value is to fish at $n = 3$.

Lastly, we can find the optimal policy at low density in the third year. Choosing to fish will mean that there is a 20% chance of the population density being medium in the next year, and a 80% chance that density stays low. In this case the expected future value will be $0.2 \times 5 + 0.8 \times 1 = 1.8$. If the decision maker chooses not to fish, the expected future value will instead be $0.1 \times 7 + 0.8 \times 5 + 0.1 \times 1 = 4.8$. As such, the optimal decision in this scenario is again not to fish. The optimal policy, as depicted in Table 4.9, is $\Pi_3(\underline{s}) = (F, F, DF))$, which gives an optimal value $V_3(\underline{s}) = (11.8, 6.9, 4.8))$. Similar calculations for the remaining years ($n = 2$ and $n = 1$) are found in Table 4.10 and Table 4.11. The policies for these years are $\Pi_2(\underline{s}) = (F, DF, DF))$ and $\Pi_1(\underline{s}) = (F, DF, DF))$, and so we note that in this stochastic decision model, the optimal policies have also converged to a stationary policy.

We see that the stationary optimal policy is the same as the deterministic case, in

F		$n+1$		
		High	Medium	Low
n	High	0.1	0.8	0.1
	Medium	0.05	0.15	0.8
	Low	0	0.2	0.8

DF		$n+1$		
		High	Medium	Low
n	High	1	0	0
	Medium	0.9	0.1	0
	Low	0.1	0.8	0.1

Table 4.7: State transition matrices for the stochastic optimisation problem, for both of the actions available to the decision maker at time n.

$n=4$	Fish		Do not fish		Optimum	Optimum
s: State	$R(s,F)$	$E(V_{t+1}(s))$	$R(s,DF)$	$E(V_{n+1}(s))$	a	$V_n(s)$
High	7	0	0	0	F	7
Medium	5	0	0	0	F	5
Low	1	0	0	0	F	1

Table 4.8: Outcomes of each choice of action during the time period $n = 4$

$n=3$	Fish		Do not fish		Optimum	Optimum
s: State	$R(s,F)$	$E(V_{n+1}(s))$	$R(s,DF)$	$E(V_{n+1}(s))$	a	$V_n(s)$
High	7	0.1×7 $+0.8\times 5$ $+0.1\times 1=4.8$	0	1×7 $+0\times 5$ $+0\times 1=7$	F	11.8
Medium	5	0.05×7 $+0.15\times 5$ $+0.8\times 1=1.9$	0	0.9×7 $+0.1\times 5$ $+0\times 1=6.8$	F	6.9
Low	1	0×7 $+0.2\times 5$ $+0.8\times 1=1.8$	0	0.1×7 $+0.8\times 5$ $+0.1\times 1=4.8$	DF	4.8

Table 4.9: Outcomes of each choice of action during the time period $n = 3$

$n=2$	Fish		Do not fish		Optimum	Optimum
s: State	$R(s,F)$	$E(V_{n+1}(s))$	$R(s,DF)$	$E(V_{n+1}(s))$	a	$V_n(s)$
High	7	0.1×11.8 $+0.8\times 6.9$ $+0.1\times 4.8=7.18$	0	1×11.8 $+0\times 6.9$ $+0\times 4.8=11.8$	F	14.18
Medium	5	0.05×11.8 $+0.15\times 6.9$ $+0.8\times 4.8=5.465$	0	0.9×11.8 $+0.1\times 6.9$ $+0\times 4.8=11.31$	DF	11.31
Low	1	0×11.8 $+0.2\times 6.9$ $+0.8\times 4.8=5.22$	0	0.1×11.8 $+0.8\times 6.9$ $+0.1\times 4.8=7.18$	DF	7.18

Table 4.10: Outcomes of each choice of action during the time period $n = 2$

$n = 1$	Fish		Do not fish		Optimum	Optimum
s: State	$R(s, F)$	$E(V_{n+1}(s))$	$R(s, DF)$	$E(V_{n+1}(s))$	a	$V_n(s)$
High	7	0.1×14.18	0	1×14.18	F	18.184
		$+0.8 \times 11.31$		$+0 \times 11.31$		
		$+0.1 \times 7.18 = 11.184$		$+0 \times 7.18 = 14.18$		
Medium	5	0.05×14.18	0	0.9×14.18	DF	13.893
		$+0.15 \times 11.31$		$+0.1 \times 11.31$		
		$+0.8 \times 7.18 = 8.1495$		$+0 \times 7.18 = 13.893$		
Low	1	0×14.18	0	0.1×14.18	DF	11.184
		$+0.2 \times 11.31$		$+0.8 \times 11.31$		
		$+0.8 \times 7.18 = 8.006$		$+0.1 \times 7.18 = 11.184$		

Table 4.11: Outcomes of each choice of action during the time period $n = 1$

that the best decision is to fish if the population density is high, and not to fish if the population density is medium or low. However, in the stochastic model it has taken longer for the time-dependent optimal policy to converge to this stationary policy than it took in the previous deterministic model.

4.3 Continuous state and action space

In practice, our toy model is not detailed enough, as exploiters will wish to know population levels with more precision than simply where it fits into a few descriptive density categories. In this section, we will use a more realistic model of fish population dynamics, taking advantage of our understanding of the solution technique derived from the use of the toy model. As in Section 4.2, we will first consider deterministic population growth (Subsection 4.2.1), and then extend the model to include stochasticity in the recruitment process (Subsection 4.2.2).

Let u be the directly measured population biomass. This gives us an exact numerical value, rather than the simple categories we saw in our toy model. Biomass values are found in a continuous, bounded range. As the population biomass is a physical quantity it must be non-negative. We define $u_{\max}$ to be the maximum instantaneous biomass the population can

take. The existence of this is physically feasible as there must be some upper limit on the possible number of fish occupying a given environment, as well as a physiological maximum size of a single fish. The population upper limit will be significantly higher than the carrying capacity K, which is the maximum population size which is sustainable at equilibrium in the long term (as opposed to the maximum possible for a short time). This upper limit is also necessary for the numerical calculation required to form a solution. Similarly, we let the exploiter precisely set their fishing effort F, the rate at which they extract biomass from the fishery. The new action space is also continuous and bounded, with allowable values being the possible fishing efforts. Efforts must only have non-negative values, as the exploiter cannot fish with negative effort. There will also be an upper limit on the allowable fishing effort $F_{\max}$, which represents the physical (due to a limited number of boats/nets/space to hold catch) or regulatory constraints on the rate at which fish can be harvested.

In this model, there is a continuous state space, representing the possible density values the fish population can take. The Bellman equation for this problem differs from a discrete state model in the manner in which the expected future value is calculated. The major functional difference that a continuous state space and action space provides is that the transition matrix Q is replaced by a probability density function p for the transitions from the current state to possible future states when an action is taken. Furthermore, the expected future value for each action is calculated by integrating (rather than summing) over all possible future states. As such, the Bellman equation becomes

$$V_n(s_n) = \max_{a \in A} \{ R(a, s_n, s'_{n+1}) + \delta \int_0^\infty p(s'_{n+1}|s_n, a) V_{n+1}(s'_{n+1}) \}. \tag{4.11}$$

4.3.1 Deterministic stock recruitment

In this more complicated model we can more accurately look at population dynamics. Specifically, we will know exactly how fishing depletes population density, and how the stock recruitment between generations occurs. The effort level F with which the fisher acts is a

constant rate of depletion over that fishing period, and so the population density decays exponentially over time, that is,

$$u_n(t) = u_n(0)\exp(-Ft). \tag{4.12}$$

Between each fishing period of length T, a new generation of the fish population will spawn. The stock recruitment relationship is given by the Beverton-Holt map:

$$u_{n+1}(0) = \frac{Ku_n(T)}{u_n(T) + (K - u_n(T))\exp(-rT)} \tag{4.13}$$

The population biomass at the beginning of the next fishing period is therefore given by the map:

$$f(u_n) = \frac{Ku_n\exp(-FT)}{u_n\exp(-FT) + (K - u_n\exp(-Ft))\exp(-rT)} \tag{4.14}$$

which acts as a transition function from state u_n at time n to state u_{n+1} at time $n+1$. We can also define an immediate payoff function, for the profit received by the decision maker in a single fishing period, based on the yield obtained that year:

$$\begin{aligned} R_n &= BF_n\int_0^T u_n(t)dt - cF_nT \\ &= Bu_n\left(1 - \exp(-F_nT)\right) - cF_nT \end{aligned} \tag{4.15}$$

Note that our definition of the immediate reward R_n is based only on the current state and action, and not on the future state. We define δ as the discount factor of future to current profits, that is, the proportion of the profit this year that we value an equal profit next year. $B > 0$ is the market price per unit fish biomass, and $C \geq 0$ is the constant cost per unit effort of fishing. Then we are able to define the problem we are faced with as finding a policy $\Pi_n : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ such that setting $F_n = \Pi_n(u_n)$ will maximise:

$$\sum_{i=0}^{N} \delta^i \left(Bu_i\left(1 - \exp(-F_iT)\right) - cF_iT\right) \tag{4.16}$$

where $u_{n+1} = f(u_n)$. Our dynamic programming equation for this deterministic model will then be:

$$V_n(u_n) = \max_{F \in A} \{R(u_n, F) + \delta V_{n+1}(u_n \exp(-F_n T))\} \tag{4.17}$$

Implementation of the model

Without loss of generality, we let the length of each fishing period be $T = 1$ year, and non-dimensionalise population levels so that the carrying capacity $K = 1$. The parameters of interest are now the discount rate, the growth rate r, and the ratio of costs to revenue from fishing, c/B. For numerical reasons, we discretise the state space S and the action space A. In order to solve this problem computationally, we perform a numerical discretisation of our continuous population density u and effort level F, using step sizes $\Delta u = 0.01$ and $\Delta F = 0.01$. We use a backwards induction algorithm to find recursively the value function for each generation of the fishing problem. The process is as follows:

1. Calculate the optimal action for each state in the final year ($n = N$) by solving:

$$V_N(u_N) = \max_{F \in A} \{R(u_n, F)\} = \max_{F \in A} \{B u_N (1 - \exp(-FT)) - c F_n T\}$$

   For each discrete value of the state variable u, perform an exhaustive search through all potential fishing efforts F, calculating the current value if this action is taken. Store the action F^* which maximises value at the present state u^* as the action prescribed by the optimal policy $\Pi_n(u^*) = F^*$.

2. Decrement time by one year, represented by single time period of the MDP.

3. Solve the recursive value problem, which takes into account the value of being in each state in the future

$$V_n(u_n) = \max_{F \in A} \{R(u_n, F) + \delta V_{n+1}(u_n \exp(-F_n T))\}$$

   As above, this is solved by exhaustively searching though all allowable fishing efforts,

and determining which value of F maximises $V_n(u)$ in each state u.

4. Save the best action as the policy for this iteration, $\Pi_n(u^*) = F^*$.

5. If the policy in this current year is the same as in the previous year (so $\Pi_n(u) = \Pi_{n+1}(u)$) for all states u, then stop, as we have obtained a stationary optimal policy. Otherwise, return to Step 2.

Figure 4.2 shows that the optimal policy obtained in each fishing period will quickly converge to the stationary value, with this convergence in fact happening with these parameter values by the penultimate year. We see that using a backwards induction solution method can in some cases obtain a stationary optimal policy very quickly, even when using a finite time horizon. Recall that a policy is a mapping from a population biomass state to a specific optimal fishing effort. A stationary optimal policy therefore tells us, under an infinite time horizon, how heavily to exploit the fish population each year, depending on the present stock level. There are several important features to note about this policy. No fishing occurs at all below a certain biomass threshold (here $u \approx 0.6$). At population levels below this point, it is not optimal to harvest at all, as the possible future value of the fish stocks will be reduced due to population depletion in the future. By not fishing at all in the current season, population levels will be replenished, and the future value of exploitation will be higher than the present one-stage reward that can be obtained by fishing. We also note that the optimal effort F increases with u. At higher population states, more fish can be harvested (giving present rewards) without heavily depleting the population and reducing the available future value of the fishery. Finally, the optimal fishing behaviour in the final year of a finite time horizon problem ($n = N$) is to fish at a higher rate than the stationary optimal policy (which is the fishing policy in previous years). In this final year there will be no future value in the fishery (with non-existent salvage value in this model), meaning it is in the monopolist's interest to heavily deplete the population. Because the total cost of fishing increases with effort, the optimal behaviour is not to simply harvest at the maximum allowable rate, meaning we still observe the positive relationship between biomass and fishing effort in the final year.

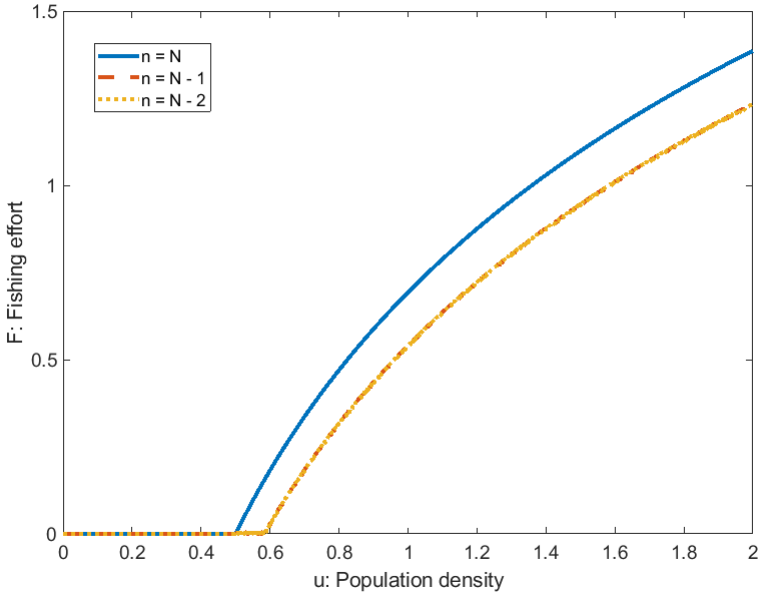
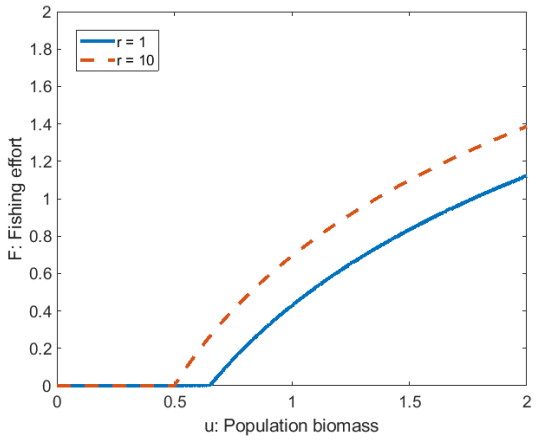
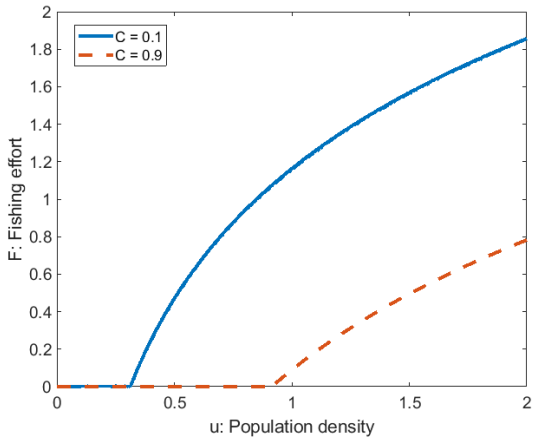


Figure 4.2: The optimal policy quickly converges such that the optimal action for each state in $N-2$ is the same as in $N-1$. We therefore have a stationary mapping in which the optimal policy is independent of time. Here $\delta = 0.95$, $C = 0.5$, $B = 1$, and $r = 2$.

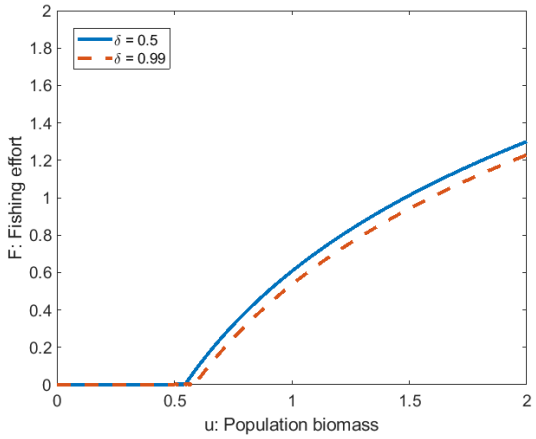
In Figure 4.3 we investigate how changing a single parameter value will change any stationary optimal policy. In each case, we see that the shape of the policy does not significantly change inside the limits of the state and action spaces, with parameter changes instead acting to change the optimal fishing effort in the same way for each state. For example, in (a) we see that increasing r causes fishing to begin occurring at lower population densities, with the optimal fishing effort therefore being higher in the high growth rate case at each density in which fishing does occur. In (b), we see that increasing the cost per unit effort of fishing causes the optimal policy to be shifted to higher levels of u (so that the same levels of exploitation happen at higher density) in the state space. (c) depicts an increase in the discount rate δ causing a slight shift of the optimal policy to higher u, and a slight reduction in its slope.



(a) Changing growth rate r: $C = 0.5, \delta = 0.95$



(b)Changing cost per unit effort C: $r = 2, \delta = 0.95$



(c)Changing discount rate δ: $r = 2, C = 0.5$

Figure 4.3: We investigate the effect of changing a single parameter on the stationary optimal policy (where convergence in time has occurred for each set of parameter values). In each row, the parameter of interest differs, with it being r in the 1st row, C in the 2nd, and δ in the 3rd.

4.3.2 Stochastic stock recruitment

Finally, we are able to complete our model to allow us to fulfil our primary objective for this chapter: finding an optimal policy (with respect to choosing from a continuous fishing effort), in a noisy continuous population. The model is the same as in Section 4.3.1, with one crucial difference. The new population density of each generation is now a random variable. Specifically, it is a $N(\mu, \sigma^2)$ normal random variable for which the mean value is the corresponding future population density in the deterministic model, and σ^2 is a fixed variance. That is, we use the stock recruitment relationship given in Equation (4.14), incorporating the exponential population decay over the fishing period and the Beverton-Holt model of reproduction, to obtain the mean population density in the next year. The standard deviation σ is the level of noise in the growth process. Formally the problem is to find a policy $\Pi_n : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ such that setting $F_n = \Pi_n(u_n)$ will maximise the *total expected value* of:

$$\sum_{n=0}^{N} \delta^n \left(B u_n \left(1 - \exp(-F_n T) \right) - c F_n T \right) \tag{4.18}$$

where $0 < \delta \leq 1, C \geq 0, B > 0, T > 0$ are constants, and u_n satisfies:

$$u_{n+1} = \begin{cases} u_{\max} & \text{if } f(u_n) + Z_n > u_{\max} \\ f(u_n) + Z_n, & \text{if } 0 < f(u_n) + Z_n < u_{\max} \\ 0, & \text{otherwise} \end{cases} \tag{4.19}$$

where the f is the deterministic mapping used in the earlier model and the Z_n are iid $N(0, \sigma^2)$ normal random variables. The expected value of being in state u_n at time n is given by:

$$V_n(u_n) = \max_{F \in A} \left\{ R(u_n, F) + \delta \int_0^{\infty} p(y|u_n, F) V_{n+1}(y) dy \right\} \tag{4.20}$$

which is the recursively defined probabilistic dynamic programming equation on which we perform the backwards induction algorithm. In this equation, $R(u_n, F)$ is the immediate

one-stage payoff function as in Section 4.3.1, when action F is taken in state u_n. This payoff is assumed to be deterministic, so that noise is confined to the stock recruitment process. The PDF for state y at time $n+1$, given that the state at time n is u_n and the individual decides to fish with effort F, is given by $p\left(y|u_n, F\right)$.

In order to solve this problem computationally, we perform a numerical discretisation of our continuous state space u and action space F, using step sizes $\Delta u = 0.01$ and $\Delta F = 0.01$. This also means that the probability density function for state transitions is also discretised. We use a backwards induction algorithm to recursively find the value function for each generation of the fishing problem. The process is as follows:

1. Calculate the optimal action for each state in the final year ($n = N$) by solving:

$$V_N(u_N) = \max_{F \in A} \left\{R\left(u_n, F\right)\right\} = \max_{F \in A} \left\{Bu_N\left(1 - \exp(-FT)\right) - cF_nT\right\}$$

   For each discrete value of the state variable u, perform an exhaustive search through all potential fishing efforts F, calculating the current value if this action is taken. Store the action F^* which maximises value at the present state u^* as the action prescribed by the optimal policy $\Pi_n(u^*) = F^*$.

2. Decrement time by one year, represented by a single time period of the Markov decision process.

3. Solve the now recursive value problem, which takes into account the value of being in each state in the future:

$$V_n(u_n) = \max_{F \in A} \left\{R\left(u_n, F\right) + \delta \int_0^\infty p\left(y|u_n, F\right) V_{n+1}(y) dy\right\}$$

   As above, this is solved by exhaustively searching though all allowable fishing efforts, and determining which value of F maximises $V_n(u)$ in each state u.

4. Save the best action as the policy for this iteration, $\Pi_n(u^*) = F^*$.

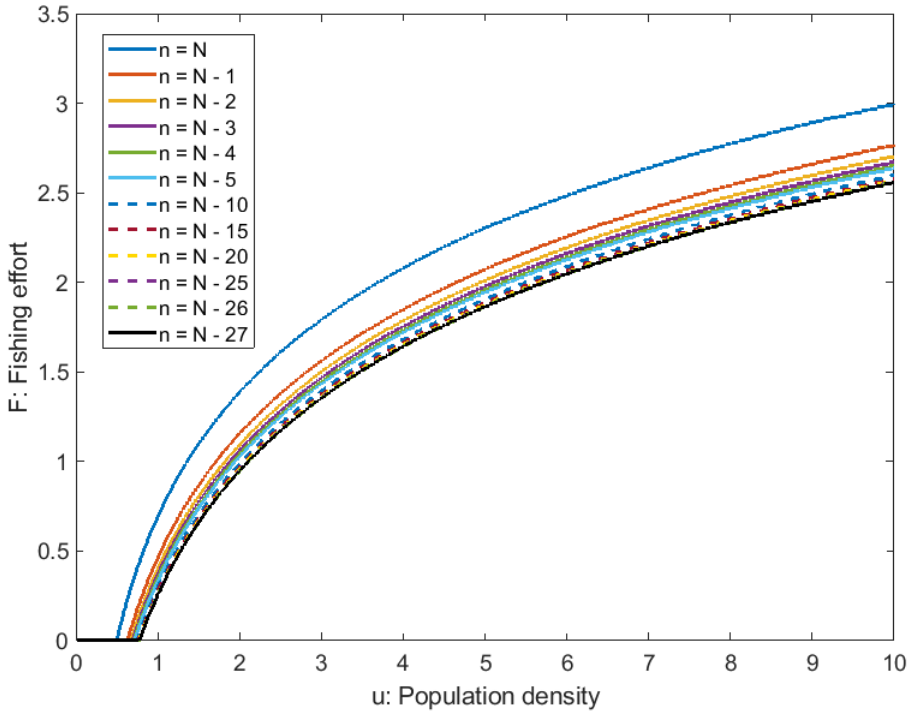
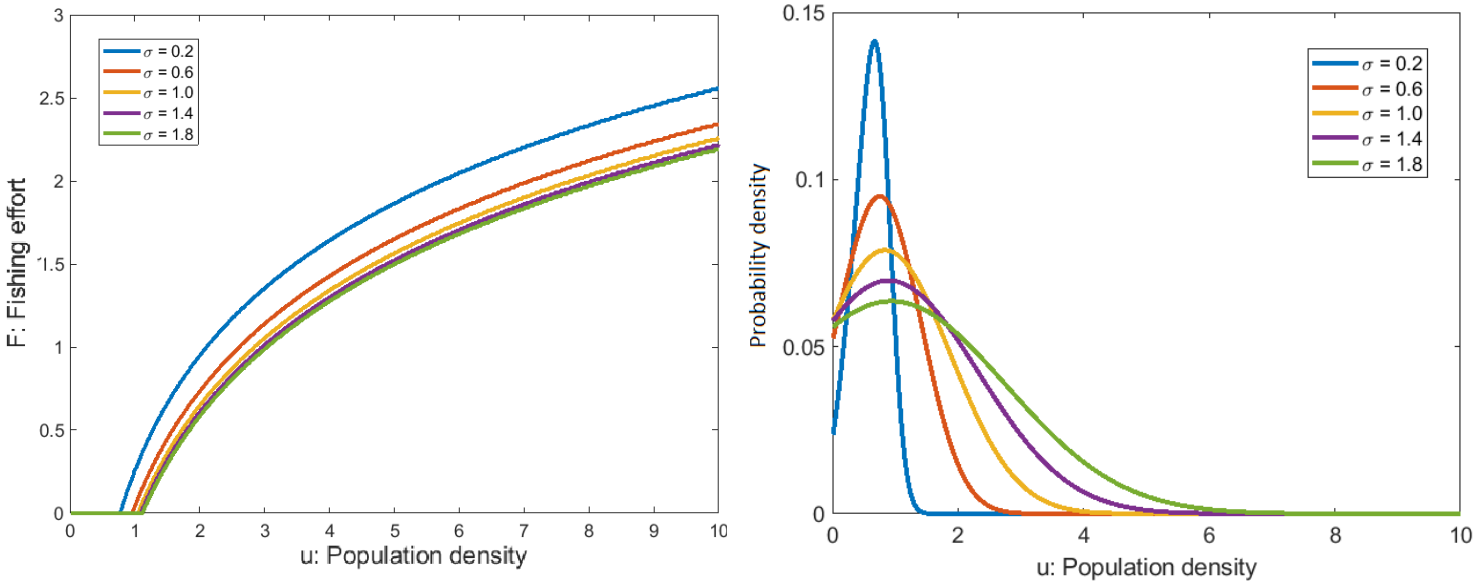


Figure 4.4: With noisy population dynamics ($\sigma = 0.2$), convergence of the optimal policy to a stationary one will still occur. Here $\delta = 0.95, r = 0.1, c = 0.5$.

5. If the policy in this current year is the same as in the previous year (so that $\Pi_n(u) = \Pi_{n+1}(u)$) for all states u, then stop, as we have obtained a stationary optimal policy. Otherwise, return to Step 2.

Using this method of solving Equation (4.20), we see in Figure 4.4 that for a fixed level of noise, policies will converge to a stationary optimal policy, as in the deterministic model. The amount of noise in the population dynamics can have a significant impact on the obtained optimal policy. Figure 4.5 (a) shows as σ increases, the optimal fishing effort at each population level will decrease. Furthermore, the minimum population size for which fishing is tenable will increase. For each value of σ in Figure 4.5 (a), the algorithm was run for n iterations until successive optimal policies had converged; that is the sum of squared differences between the specified effort levels at each state in iteration n and iteration $n-1$ was sufficiently small.



(a) (b)

Figure 4.5: (a) shows that increasing the stock-recruitment noise σ will cause the optimal fishing effort to decrease with population size. (b) gives the stationary distribution of the population, conditional upon non-extinction, given by the dominant eigenvector of the matrix M_σ, under fishing according to a stationary optimal policy. For each policy, $\delta = 0.95, r = 0.1$, $B = 1$ and $c = 0.5$.

When a stationary policy has been obtained, the transition matrix (where our continuous state space S has been discretised) Q_σ will be constant for each successive iteration. Note that the transition matrix depends on the stochastic stock recruitment process, and so the probability of transitioning between specific states depends on the specific value of the population variance σ. For a stationary policy, as the specified optimal action at every population density will be the same regardless of time, it follows that the probability of transitioning from any one state to another will also be constant, and so Q_σ is time invariant.

Q_σ is a stochastic matrix, and so has a dominant eigenvalue of 1. We note that this transition matrix includes transitions to negative population densities, which are non-viable in our fishery model. Indeed, any transition to a population state $u \leq 0$ represents the extinction of the fish population, with no ability of future population recovery, as described in (4.19). We therefore define M_σ as a new transition matrix, which is the restriction of Q_σ to the

viable population states $0 < u < u_{\max}$. This matrix M_σ is not a stochastic matrix: its dominant eigenvalue is less than 1. The dominant eigenvalue λ is the probability of survival, or the probability that the fish population will not go extinct in a single time period, given that the optimal policy is followed. In each time period, there is a probability that the fish population transitions to an extinction state. These extinction states are represented by the entries of Q_σ that are not in M_σ, while the states in which the population survives are in M_σ.

The dominant eigenvector of the restricted transition matrix M_σ gives a stationary population distribution of viable states, conditional on the non-extinction of the population, that will result when the stationary policy is applied over time to a population with the corresponding noise level. In Figure 4.5 (b), we observe the eigenvectors corresponding to the dominant eigenvalue of the restricted transition matrix, for a range of recruitment noise. Each eigenvector is plotted against the state variable u, and represents the stationary distribution of the population that will occur under a stationary optimal policy. We observe that the peak values of these stationary distributions move to the right in the state space as noise increases, such that the modal population density, conditional upon non-extinction, under optimal exploitation is larger with greater noise. We also note that increased noise causes a larger spread in the stationary population distribution, with an increased probability of the population going extinct (where $u = 0$) in a single fishing period.

The dominant eigenvalues λ_1 of the restricted transition matrix M_σ give the probability that extinction will not occur in the next fishing period. If λ_1 is much less than one, then a large proportion of the PDF for each transition will be (before restriction of the PDF to the allowed state space) for population levels below 0, meaning that extinction is likely. Calculating the dominant eigenvalue therefore gives us a mechanism to investigate the likelihood of the fish population going extinct under certain conditions. For Figure 4.6, with a range of noise levels, we have calculated both the optimal policy under that level of noise, and the transition matrix that corresponds to using that policy in that noisy environment. Figure

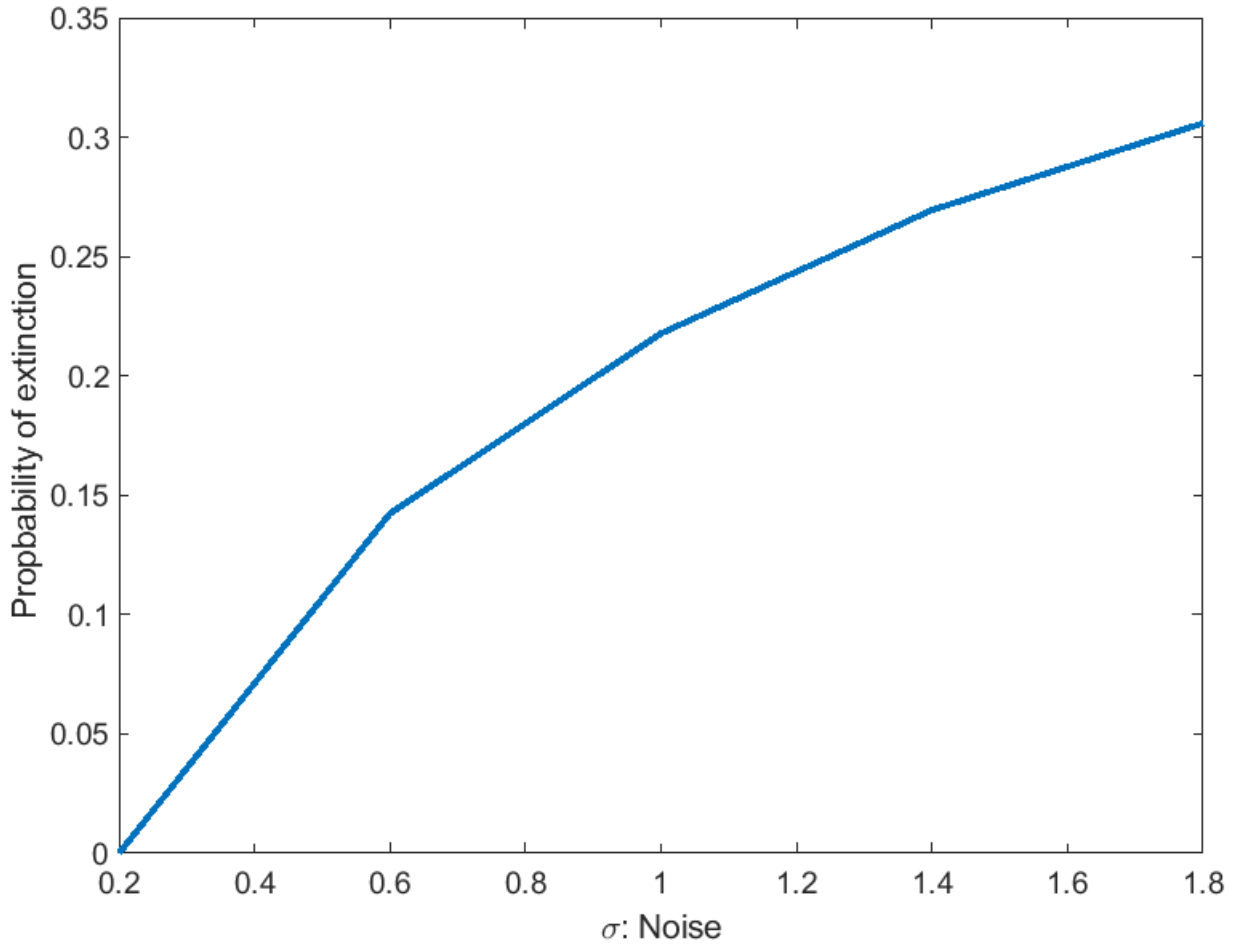


Figure 4.6: The probability of extinction $(1 - \lambda_1)$ in a single time period, when using both the optimal policy and the deterministic policy, for a range of noise levels. Here $\delta = 0.95, r = 0.1, c = 0.5$.

4.6 shows the probability of extinction for each level of noise when using that optimal policy. We see that as the variance in the stock-recruitment relationship increases, the probability that a population will go extinct in a single generation also increases.

4.4 Discussion

We offer an explanation of the mechanisms behind the effect of parameter changes on optimal policies. By decreasing the discount rate δ, we increasingly value current profits more than future profits. As such, less effort will be made to preserve the state of the fishery by

maintaining population levels, because doing so will be economically less viable according to the decision maker's priorities and motives. Because of this, we expect the optimal policy with a lower discount rate to be one in which there is a similar or greater level of fishing effort occurring at each population density. If the normalised cost per unit effort level is decreased, then it will be comparatively cheaper for the decision maker to fish, with single-stage profit maximisation occurring at higher levels of fishing effort. Consequently, low costs will be associated with taking advantage of immediate profits, with optimal policies being for heavy levels of exploitation to occur, even at lower population densities. With a high growth rate r, we expect the population levels to quickly replenish to near the carrying capacity after it has been exploited. High r will therefore be associated with a reasonable amount of resilience to heavy rates of harvesting, allowing optimal fishing behaviour to occur at higher intensity levels. This influence of parameter values on fishing policies can be expected to hold for both the single player MDP model, and the two-player Markov game model.

Our chosen algorithm, backwards induction, is usually used for finite time horizon problems rather than infinite horizon ones. However, as we know that the one-stage profits are bounded, and that this means that a stationary optimal policy must exist, we can iterate through stages of the MDP until convergence occurs. Essentially, we set the length of the finite time horizon to be sufficiently long enough that we obtain a solution to the infinite time horizon problem. It is likely that alternative solution methods such as value iteration or policy iteration could be used to find an equivalent solution, however, our dynamic program was sufficiently tractable that this was not necessary.

Although our algorithm for the single fisher model is guaranteed to provide us with an optimal stationary policy, it is not necessarily a unique one. If it any stage of the algorithm there are two or more states with the same value, then the algorithm may in the next iteration choose actions that would lead to either state without discrimination. We do however note that while there may be more than one optimal policy, the actual value of being in each

state in each of these policies would be identical, with the specified actions to take being what differs between the policies.

An interesting feature of the stochastic continuous model is the effect of changing population noise levels on the optimal policy. We would expect, and indeed observed in Figure 4.5, that any increase in the noisiness of the population dynamics will cause the fish population to be more likely to go extinct under a given level of fishing effort. As such, the decision maker would have a greater incentive to preserve population levels, as extinction will be more likely to occur with the impact of fishing. This will effectively act to disincentivise profiting in the short term over possible profits in the future, equivalent to reduced discounting of future profits. That is, it is optimal for the decision maker to lower their fishing effort in a more noisy environment, in order to protect their possible future gains from the fishery.

One interesting possible adaptation to this model would be to add a multiplicative noise term. Currently, the noise in the population dynamics is independent of the population level. Adding density dependent noise would allow us to capture an increased range of behaviour, with the variability of population growth changing with population density. For example, it would mean increasing the stochasticity of the stock recruitment relationship, as with higher numbers of fecund adult fish we could expect greater variability in the number of new fish spawning.

Chapter 5

Multi-agent exploitation of a stochastic fishery: a Markov game model

5.1 Introduction

Markov decision processes are an effective tool for solving stochastic decision making problems when there is only a single decision maker. However, they are not suitable to use when there are multiple decision makers able to influence outcomes. Instead, these types of scenarios can be modelled as stochastic games, which are generalisations of MDPs to two or more players in which the rewards and future state are affected by the actions of all players (Hu et al., 1998). Also known as Markov games, stochastic games are equivalently a generalisation of repeated games (as seen in Chapter 3) to allow for probabilistic transitions between states. We will restrict our investigation to a stochastic game with two players, though games with a greater number of players will operate conceptually similarly.

5.1.1 Stochastic games

A two-player stochastic game consists of (Littman, 1994)

- A set $I = \{1, 2\}$ of players
- A state space S
- An action set $A^{(i)}$ for each player $i \in I$
- For each combination of states s, s' and actions $a^{(i)}, a^{(j)}$, the transition probability $Q(s, s', a^{(i)}, a^{(j)})$; this is the probability that the choice of action $a^{(i)}$ by player i in state s will lead to state s', given that player j chooses action $a^{(j)}$
- $R^{(i)}(s, s', a^{(i)}, a^{(j)})$ is the immediate reward (or one-stage payoff) for player i when a transition from state s to state s' occurs when action $a^{(i)}$ has been taken by player i, and player j chooses action $a^{(j)}$.
- δ is the discount rate of future rewards

In solving a stochastic game, we aim to find an optimal policy $\Pi^{(i)}$ for each player $i \in I$. These policies are the same as the solution to a Markov decision process, in that they are a mapping from the state space S to the action space A, and represent the best action $a_n^{(i)} = \Pi_n^{(i)}(s_n)$ to take at time n in order to maximise their own expected future profit, or value, when the fish population is in state s_n. The value function for player 1 $V_n^{(1)}(s_n)$ is the expected discounted reward for player 1 from time n until the end of the time horizon, given that the state is s_n at time n, and is given by

$$V_n^{(1)}(s_n) = R^1(s_n, s', a_n^{(1)}, a_n^{(2)}) + \delta \sum_{s' \in S} Q(s + n, s', a^{(i)}, a^{(j)}) V_{n+1}^{(1)}(s_n) ds \tag{5.1}$$

Analogously to the Bellman equation (Equation 4.4), we can define a recursively solvable dynamic programming equation that allows us to obtain the optimal action in each instance (Littman, 1994). A solution to a Markov game must be simultaneously optimal for each player; we will look for a Nash equilibrium which here will be a pair of policies $\Pi^{(1)}$ and $\Pi^{(2)}$ such that when player 1 is playing $\Pi^{(1)}$, player 2's optimal policy is $\Pi^{(2)}$, and vice versa.

5.1.2 Solution methods for stochastic games

While all finite Markov decision processes have a (not necessarily unique) stationary optimal policy as their solution, this is not the case for Markov games. Indeed, we know that even for deterministic matrix games, there will not necessarily be any pure Nash equilibrium, as seen in the well-known matching pennies game (Bowling and Veloso, 2000). Stochastic games with a finite number of players, and finite state and action spaces always have a Nash equilibrium (Shapley, 1953; Filar and Vrieze, 2012). However, this Nash equilibrium is in general not deterministic or unique (Hu et al., 1998), and there indeed exist games which have no deterministic equilibrium policies (Zinkevich et al., 2006). The majority of work in solving stochastic games has focussed on zero-sum games, which have a unique Nash equilibrium (Bowling and Veloso, 2000) rather than the general-sum games we consider, for which there may be multiple Nash equilibria. Value iteration has been used to obtain solutions in two-player zero-sum games (Shapley, 1953) and two-player general-sum games (Filar and Vrieze, 2012). Policy iteration methods have been used (Vrieze, 1987; Littman, 1994; Hu and Wellman, 2003), as well as reinforcement learning approaches (Hu et al., 1998) which do not explicitly model the environment of the game, but attempt to find an equilibrium policy for a single agent by focussing on their behaviour. (Bowling and Veloso, 2000). These reinforcement learning techniques have been applied to both zero-sum (Littman, 1994; Hu et al., 1998) and general-sum (Hu and Wellman, 2003) stochastic games. Individual solution methods for stochastic games do not always converge to optimal policies (Bowling and Veloso, 2000).

There has been some consideration of other solution types for stochastic games, including cyclic equilibria (Zinkevich et al., 2006), solutions which are sequences of stationary policies. Value iteration algorithms are able to identify cyclic equilibria for games in which they fail to reconstruct optimal policies (from state-action values) to find stationary equilibria (Zinkevich et al., 2006). Our analysis will focus on finding stationary Nash equilibria, such that outcomes from duopolistic exploitation of the fishery can be compared to the monopolistic

policies obtained from an MDP, and to those from the deterministic equilibrium population game in Chapter 2. We will mathematically model a fishery such that the structure of pay-offs and the transition probabilities are known to all players. This will allow us to apply Bellman-like value iteration techniques to solve a non-linear dynamic programming equation (Filar and Vrieze, 2012), in order to find Nash equilibria for our stochastic game.

5.2 Duopolistic exploitation: a stochastic game

In the real world, few fisheries are exploited by a sole agent with exclusive access. Instead, multiple independent agents with goals which may be at odds with each other will each vie for the rewards available from the fishery. In the previous chapter we detailed a Markov decision process methodology to determine the optimal fishing policy when a single agent harvests a stochastic fish population in a monopolistic scenario. This optimal policy was a stationary deterministic mapping from the state space, representing the fish population biomass, to the action space, representing the rate of fishing effort that the monopolist exploits the population with. An optimal policy therefore details the ideal fishing effort for each possible population level. We will now extend this methodology to allow for two competing agents. The simplest possible extension to our existing strategy is to add a second agent with fixed fishing behaviour. In this scenario, the second agent's fishing effort could be treated as a constant part of the environment and incorporated into our model of population dynamics. However, we wish to allow for multiple independent agents, who can each alter their harvesting behaviour in responses to the others' actions, and the state of the population. We will therefore model the non-cooperative interaction between the two fishers as a stochastic game.

5.2.1 Population modelling

Consider a symmetric stochastic game in which two identical fishers ($I = \{1, 2\}$) exploit the same fishery. Each fisher, or player, has identical costs of fishing, and identical preferences

for current and future rewards. Let the state space S and the action space A be as in the monopolist model in Chapter 4, such that fishers can choose from a continuous range of fishing efforts $0 \leq F \leq F_{\max}$, and the fish population can have densities in the range $0 \leq u \leq u_{\max}$. With two fishers, the total rate at which the population is depleted due to fishing at time n is the sum of each player's individual fishing efforts $F_n = F_n^{(1)} + F_n^{(1)}$ and so

$$f(u_n, F_n^{(1)}, F_n^{(2)}) = \frac{K u_n \exp(-(F_n^{(1)} + F_n^{(2)})T)}{u_n \exp(-(F_n^{(1)} + F_n^{(2)})T) + \Big(K - u_n \exp(-(F_n^{(1)} + F_n^{(2)})T)\Big) \exp(-rT)}. \tag{5.2}$$

The fish population grows according to a stochastic Beverton-Holt process

$$u_{n+1} = \begin{cases} u_{\max} & \text{if } f(u_n) + Z_n > u_{\max} \\ f(u_n, F_n^{(1)}, F_n^{(2)}) + Z_n, & \text{if } 0 < f(u_n, F_n^{(i)}, F_n^{(j)}) + Z_n \leq u_{\max} \\ 0, & \text{otherwise} \end{cases} \tag{5.3}$$

where the Z_n are iid $N(0, \sigma^2)$ normal random variables. Like the noisy monopolist model from the previous chapter, the Beverton-Holt map gives us the expected new population density; that is, the new population density at generation $n+1$ is a normal random variable with mean $f(u_n, F_n^{(1)}, F_n^{(2)})$ and fixed standard deviation σ, which is the parameter representing the amount of stochasticity in the stock recruitment relationship.

Players will receive a proportion of the total yield according to the ratio between their own effort and the total effort. Immediate rewards provide profit to a player based on the amount of biomass they harvest in that time period, and so the one-stage reward for player i for fishing with effort $F_n^{(i)}$ at stage n is

$$R^{(i)}(u_n, u_{n+1}, F_n^{(i)}, F_n^{(j)}) = \frac{F_n^{(i)}}{F_n^{(i)} + F_n^{(j)}} B u_n \left(1 - \exp\left(-(F_n^{(i)} + F_n^{(j)})T\right)\right) - c F_n^{(i)} \quad \text{for } i = 1, 2 \; j \neq i. \tag{5.4}$$

We note that in this fishery model the one-stage reward has no dependence on the state u_{n+1} being transitioned to; as such arguments of u_{n+1} will be omitted from the reward function in the remainder of the chapter.

As in the monopolist model, each player in this stochastic game is motivated to maximise the sum of their own (discounted) profit over an infinite time horizon. Player i's goal is then

$$\max_{F_n^{(i)} \in [0, F_{\max}]} \{\sum_{n=0}^{N} \delta^n R^{(i)}(u_n, F_n^{(i)}, F_n^{(j)})\} \tag{5.5}$$

A player maximises their expected total reward by playing a best response to their opponent's strategy, such that no unilateral deviation from this strategy will improve their own payoff at any stage of the stochastic game. By our assumption of symmetry in the information structure and preferences of the fishers, the optimal policies will be identical for each player. We are then able to construct a dynamic programming equation analogous to Equation 3.16 for the value available to player 1 in our stochastic game. As there is a continuous state space and continuous action space in this problem, the transition matrix Q is replaced in the formulation of the problem by a probability density function p, and we integrate over the possible continuous range of efforts when optimising action:

$$V_n^{(1)}(s_n) = R(s_n, a_n^{(1)}, a_n^{(2)}) + \delta \int_0^{\infty} p(s'|s_n, a_n^{(1)}, a_n^{(2)}) V_{n+1}^{(1)}(s_n) ds \tag{5.6}$$

5.2.2 Multi-agent value iteration for stochastic games

We implement an adapted process of backwards iteration as a means of obtaining simultaneously optimal policies $\Pi_n(u)$ for each player. Identically to the single agent Markov decision process, policies are developed by iterating backwards through time, year by year, beginning with a final time period in which agents do not value future profits at all. We then iterate between players in this last time period until the policies within this year converges to obtain $\Pi_n^{(1)}$ and $\Pi_n^{(2)}$. How does the process of iterating policies within a time period occur?

An initial guess for player 2's strategy is made. We then optimise player 1's behaviour by numerically determining the best fishing effort to use for each possible population state, with respect to player 2's existing policy and the effect it has on the fish population. Spline interpolation of the value function is used ("interp1" in MATLAB) to estimate the value between the discretised population states, with "fmindbnd" used to determine the fishing effort which numerically maximises the player's profit.

In response, we optimise player 2's fishing effort given player 1's newly determined policy. We repeat this process of alternating between each player and taking their opponent's new actions into account, until the behaviour of the players does not change between iterations. At this point, players' policies have converged to a fixed policy for this time period, and we deem them to be optimal in the Nash equilibrium sense (for this single stage-game). That is, each player is playing a best response to the other's possible choices of action, and neither player is able to unilaterally increase payoffs by deviating from their derived policy.

We iterate backwards through prior years in the same manner as in the MDP algorithm, following the process of iterating between players within each year until there is convergence to policies which are mutually best responses to each other. When the harvesting policy for a given player does not change from year to year (meaning the optimal fishing behaviour does not change over time) we have convergence to a stationary optimal policy. At each period with an infinite time horizon players following this optimal policy will be using best responses to each other's strategy. In this chapter, we will refer to the stationary optimal solutions of the two-player game as the Nash equilibrium. As these solutions are contingent only on the current time, and cannot take into account the history of the game, they are open-loop equilibria. For computational purposes, we discretise the state space, action space, and PDF of the problem, using mesh size $\Delta u = 0.01$ and $\Delta F = 0.01$. When solving the backwards induction equation, we numerically integrate over the action space.

5.3 Open-loop Nash equilibria

5.3.1 Deterministic stock recruitment

Following on from the single player MDP model, a natural place to start is to consider a deterministic version of this two-player Markov game (where there is no stochasticity in the stock-recruitment relationship). With $\sigma = 0$, the transitions between population states in each generation will be a deterministic function (Equation (3.18)) of current population biomass and both players' fishing efforts, with the transition probabilities therefore being given by a Dirac delta function. The recursive dynamic programming equation for player i in this deterministic game is

$$V_n^{(i)}(u_n) = R(u_n, F_n^{(i)}, F_n^{(j)}) + \delta V_{n+1}^{(1)}(f(u_n, F_n^{(i)}, F_n^{(j)})) \quad \text{for } i = 1, 2 \; j \neq i \tag{5.7}$$

To solve this equation computationally, we discretise the continuous state space u using a step size $\Delta u = 0.01$. For the deterministic version of this two-player game, we solve a continuous optimisation problem for the fishing effort F using spline interpolation. The initial conditions for each players' policy is set at half the monopolist policy, calculated as the optimal one-player fishing policy optimising Equation (4.18) in Chapter 4. Figure 5.1 (a) and (c) show the optimal Nash equilibria policies (which are the same for both players due to the symmetry of information and preferences) obtained by our algorithm for the final four time periods of this deterministic multi-stage game. Recall that due to the symmetric nature of the two-player game, the optimal policies for player 1 and player 2 are equal for each iteration. In (a), convergence occurs quickly, with the policies in $n = N - 1$ and $n = N - 2$ being the same for each value of the state variable. In 5.1 (b), we note that for this number of decision epochs, policies have not yet converged to a stationary optimal policy. If the derived policies are calculated and plotted for a larger number of stock generations, we find that they do not converge in the deterministic case. Instead, there is a large amount of numerical noise observed in the numerically obtained policies. We believe this noise occurs due

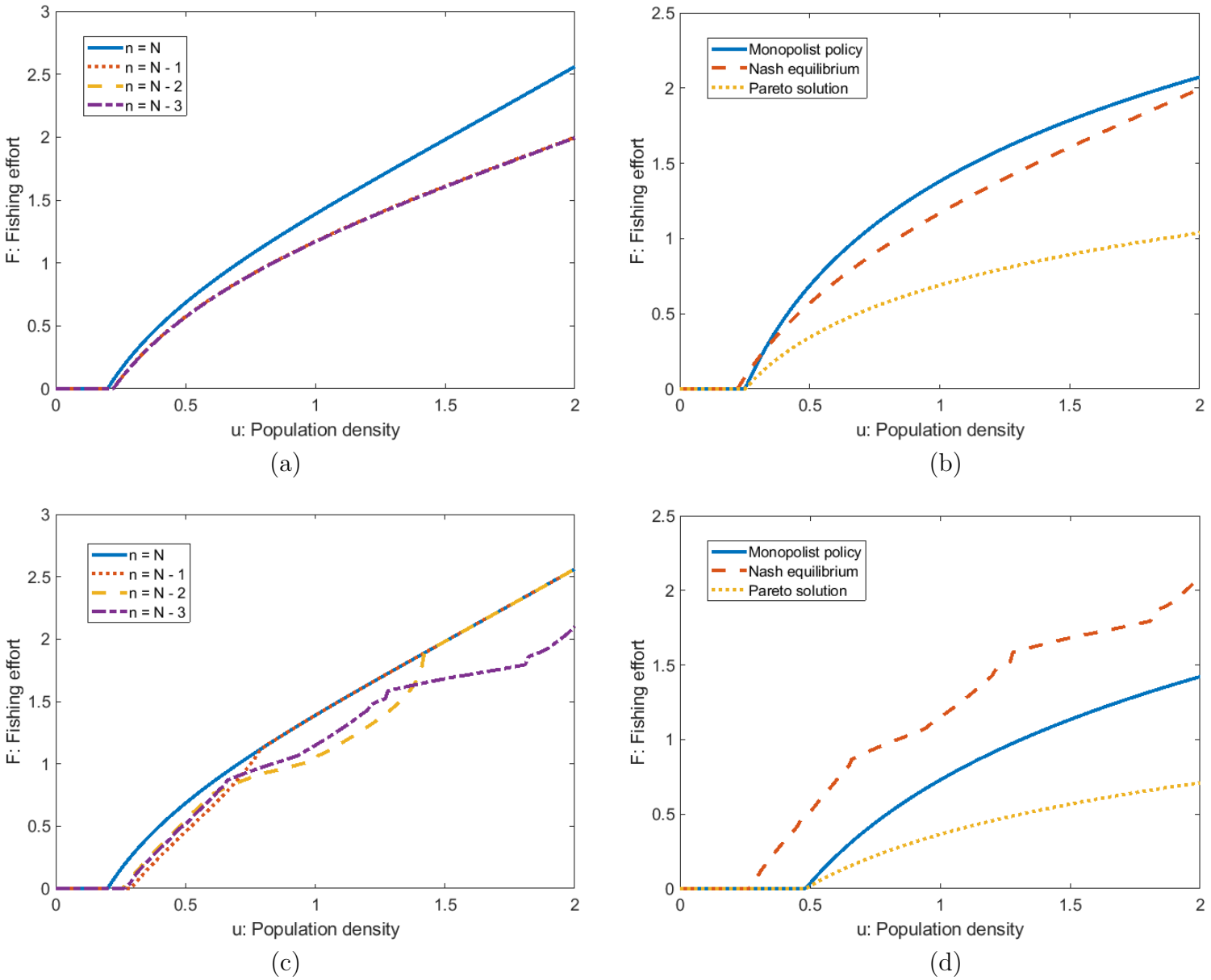


Figure 5.1: (a) and (c) show the calculated optimal policies for both players for the final four time periods in a two-player game with no noise. In (a), there is a high population growth rate ($r = 4$), and the optimal policy is reached in the penultimate fishing period. In (c), $r = 1$, and convergence over time to an optimal policy does not happen. Policies give the best choice of fishing effort F at each value of the state variable u, in terms of maximising the sum of current and expected future rewards. (b) and (d) depict each type of fishing policy in the period $n = N - 3$, for the growth rates $r = 4$ and $r = 1$ respectively. The Nash equilibrium is the optimal two-player policy as shown in (a) and (c). The monopolist policy is the best single player policy, as calculated in Chapter 4. The Pareto solution is half the monopolist fishing effort at each state. In each case $C = 0.25$ and $\delta = 0.9$. Convergence to these policies within each fishing period n has occurred.

to numerical effects such as rounding having an increased influence on the decision problem, with these initially small numerical differences being amplified as the future value component is increasingly prioritised in making optimal fishing decisions. Using a finer numerical mesh for the population state space did not allow convergence to occur, and instead amplified the effects of numerical errors. We conclude that this method of solution is not numerically stable for the deterministic game, as it is not robust to all feasible choices of parameter values.

There is still some useful information that can be gleaned from this result, as convergence can occur with high growth rates. As in the monopolist policy derived in Chapter 4, there is a minimum level of population biomass for which fishing is viable. The fishing efforts specified by the Nash equilibrium policy are also lower at each state value than the policy for the final year ($n = N$). Recall that in the final year we simply have a one-stage optimisation problem, where no individual fisher values the future state of the fishery at all. We are therefore not surprised to see that the optimal fishing effort in the final year is higher than in previous years.

We can also compare the individual Nash policy (at $n = N - 3$) to the 'socially optimal' harvesting that happens when both players fish at half the monopolist policy in Figure 5.1 (c). Here we see that at the fishing effort chosen when using the individual Nash equilibrium policy is higher, at each population state, than the socially optimal behaviour. The minimum population for which fishing is viable is approximately the same for the Nash policy and the socially optimal Pareto solution. 5.1 (d) shows that for each population state u, the individual fishing effort at the Nash equilibrium is higher than the fishing effort for the Pareto solution, which is the same relationship between efforts and solutions of the game that was seen in the prisoner's dilemma model in Chapter 3.

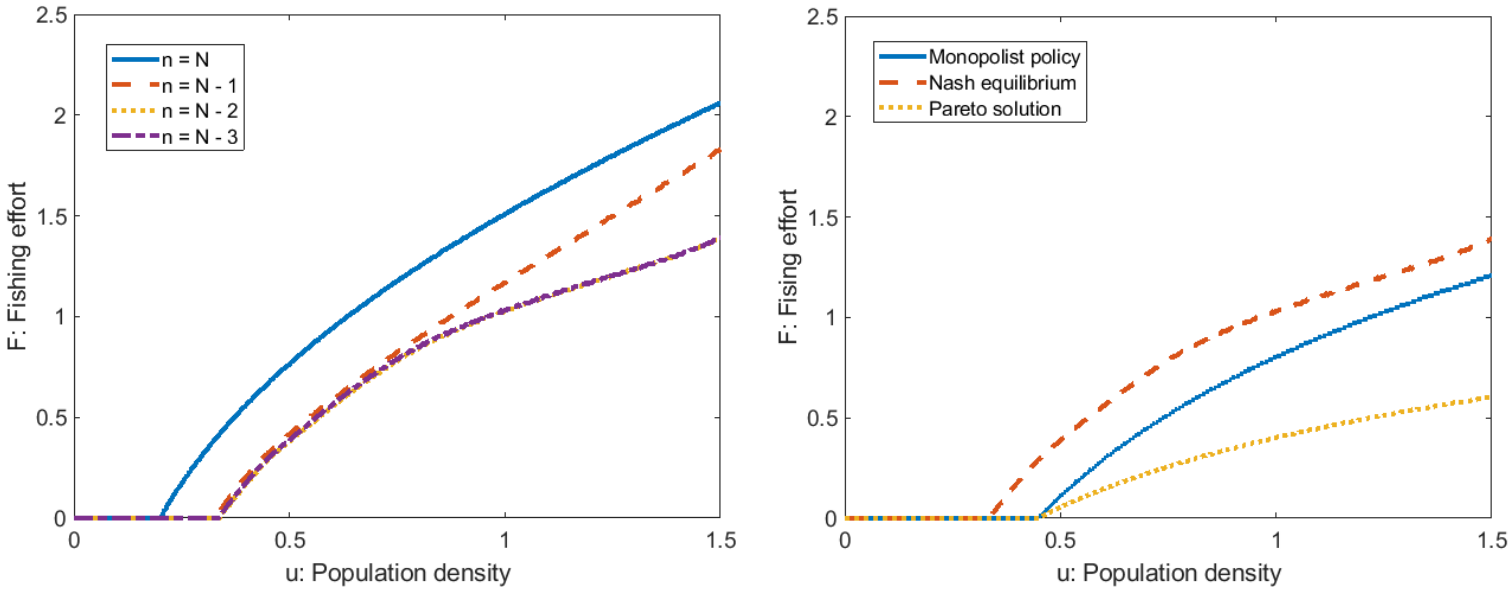
5.3.2 Stochastic stock recruitment

By adding a small amount of noise to the stock-recruitment relationship, we attempt to gain an understanding of how the presence of two decision makers exploiting the population will affect fishing under our model of population dynamics. In order to obtain a baseline comparison to the single player deterministic model, we ideally want a sufficiently small level of noise that population dynamics are as close to deterministic as possible, while the "smoothing" effect of a small amount of randomness on state transitions reduces the impact of numerical error. Essentially, we want to know if our method will still allow for convergence to an optimal policy at low levels of noise.

Unlike in the deterministic two-player game, where we solved for a continuous fishing effort F, here we numerically discretise both the state space and the action space. In each value calculation, we then perform an exhaustive search through all possible actions in order to find the specific effort which maximised the value function in each state.

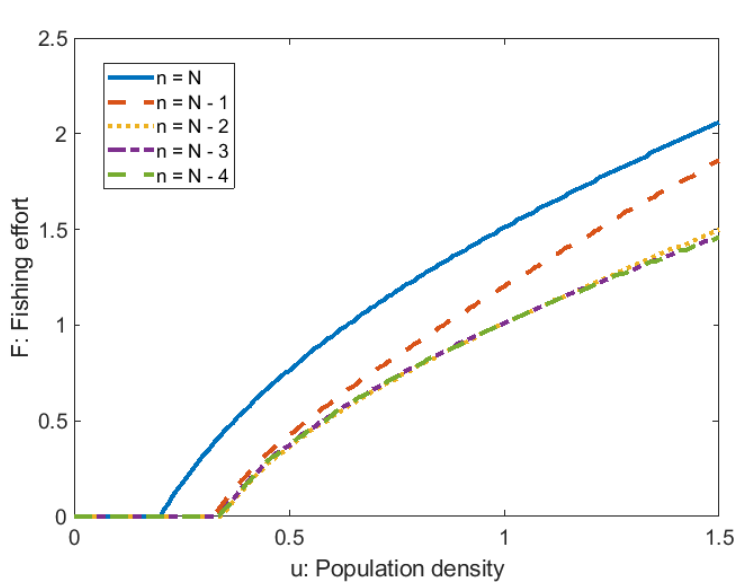
In Figure 5.2 (a) and (c), we see that convergence will occur through time to an optimal policy when there is noisy stock-recruitment. In (a), the stock-recruitment noise level is $\sigma = 0.02$ (or 2% of the population carrying capacity), and in (c) it is $\sigma = 0.1$. For the lower noise level, it takes four fishing periods for the optimal fishing policy to converge to a stationary policy. With higher noise it takes five periods. These stationary Nash equilibria have some consistent features regardless of the strength of the noise. There is a minimum population biomass for fishing to be tenable; below this population level it is optimal not to fish at all. Above the minimum level of fishing, the optimal fishing effort increases with the density of the fish population.

The stationary policy in each case is less than the optimal fishing behaviour in the final period $n = N$. This final period represents the last chance to fish in a finite time horizon problem, meaning that there is no value in maintaining population levels into the future.

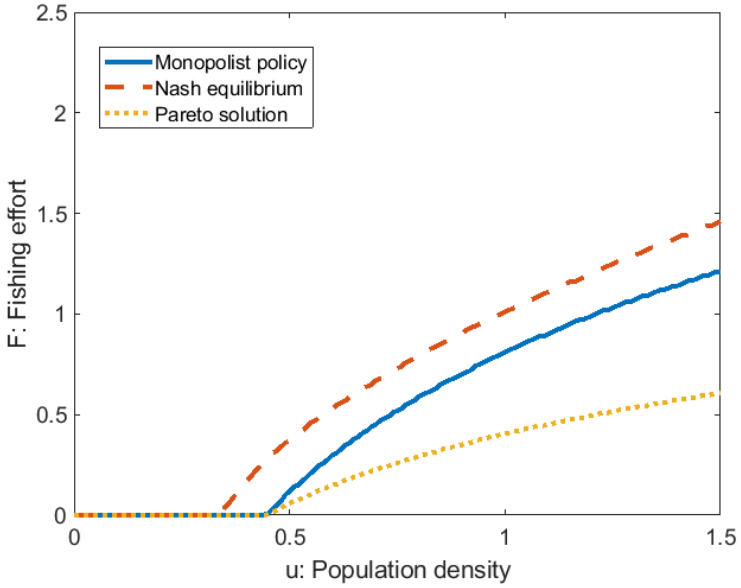


(a)

(b)



(c)



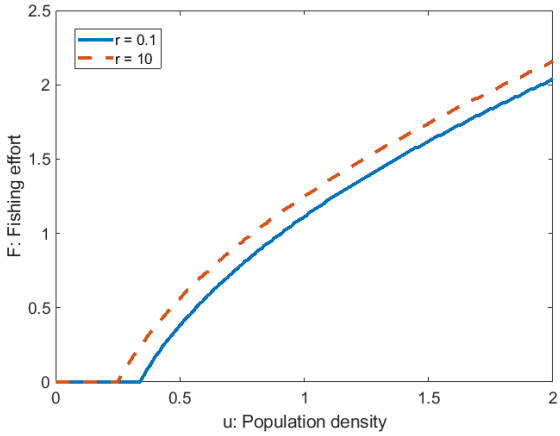
(d)

Figure 5.2: There is convergence to a stationary Nash equilibrium policy for different values of the recruitment noise σ. (a) and (c) show the optimal policy for each iteration until they converge over time. Noise levels for these figures are $\sigma = 0.02$ and $\sigma = 0.1$, respectively. (b) and (d) show each type of solution we have considered, for $\sigma = 0.02$ and $\sigma = 0.1$. The solution types are the monopolist equilibrium, as calculated in Chapter 4, the Nash equilibrium, which is the stationary optimal policy from (a) and (c), and the Pareto policy, which is half the monopolist policy at each state. $r = 2$, $C = 0.2$, and $\delta = 0.9$ for each calculation.

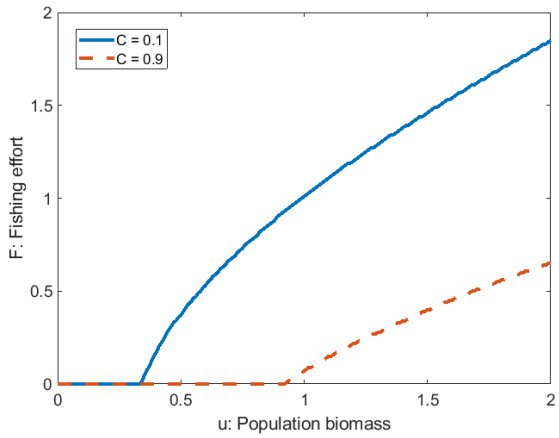
Only the costs of fishing constrain the level of exploitation. In previous periods, or with an indefinite number of fishing periods in the future, fishers' value having stock available to harvest in the future, and to allow population levels to replenish so further exploitation can occur. This means that the Nash equilibrium policies prescribe lower rates of fishing effort for each state than the policy for the final year.

As in the monopolist model, we are interested in how optimal policies fundamentally change with a change in parameters of the bioeconomic model. Each component of Figure 5.3 shows how optimal policies change with a single altered parameter. The parameters considered are c, the cost of fishing per unit effort, in Figure 5.3 (a), r, the intrinsic population growth rate, in Figure 5.3 (b), and the discount rate δ in Figure 5.3 (c). We observe that increasing the growth rate r results in slightly higher rates of optimal fishing effort F at each value of the state variable u. Decreasing the cost c has a large effect on reducing the optimal fishing effort in each state, with the minimum biomass required for fishing to occur being close to the carrying capacity when $c = 0.9$. Lastly, increasing δ results in a slight decrease in the optimal fishing rate at each state.

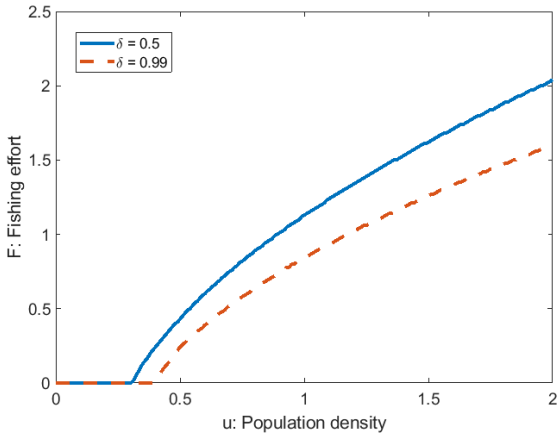
Our major purpose of investigation with this model is to see how noise in stock-recruitment impacts fishing behaviour. Figure 5.4 (a) and (c) shows how stationary optimal policies change with population noise, for populations with different growth rates. (b) and (d) show the stationary state distributions for each noise level. As in Chapter 4, the stationary state distributions, conditional upon non-extinction, are in each case the dominant eigenvector of M_σ, (the transition matrix restricted to the allowed state space) when a stationary policy is applied. In (a) and (b) there is a low population growth rate $r = 0.1$. (c) and (d) have the high growth rate $r = 10$. In (c), we see that a high growth rate, the strength of noise in recruitment has no effect on the Nash equilibrium policy. Noise levels ranging from $\sigma = 0.01$ (1% of the carrying capacity) to $\sigma = 0.3$ (30% of the carrying capacity) were considered. For each noise level, the same optimal fishing policy was obtained. That is, the best choice of



(a) Changing growth rate r: $c = 0.25, \delta = 0.9, B = 1$



(b)Changing cost per unit effort c: $r = 1, \delta = 0.9, B = 1$



(c)Changing discount rate δ: $r = 1, c = 0.25, B = 1$

Figure 5.3: We investigate the effect of changing a single parameter on the stationary optimal policy. In each row, the parameter of interest differs, with it being r in the 1st row, c in the 2nd, and δ in the 3rd.

fishing effort (to maximise total expected discounted value) is the same for each value of σ, across the entire state space. These results exhibit the same features previously seen in such policies. No fishing will be carried out when the population biomass is below $u \approx 0.5$, with the optimal fishing effort increasing with biomass above this minimum value. (d) depicts the stationary state distributions when each policy is followed. These distributions are symmetric around their mean value, and the mean population biomass is the same ($u \approx K$) for each value of σ. As population noise increases, the distributions become increasingly spread out.

Figure 5.4 (a) and (d) show results for populations with slow growth. In (a), we observe that for noise levels between $\sigma = 0.01$ and $\sigma = 0.1$, the optimal fishing policy is the same, in terms of the same optimal action being associated with each value of the state variable u. As the population noise increases even further (from $\sigma = 0.1$ to $\sigma = 0.3$), the Nash equilibrium policy changes. In this range of parameter values, an increase in σ corresponds to a lower optimal fishing effort for each population state. This is a similar result to that seen for monopolist fishing behaviour in Chapter 4. However, in this case the difference between Nash equilibria for the different noise levels is small. That is, the optimal fishing effort for each state when the noise is 30% of the carrying capacity is only slightly less than the optimal fishing effort when the noise is 10% of the carrying capacity. Figure 5.4 (b) shows the stationary state distribution corresponding to each stationary Nash equilibria being followed. As σ increases, the modal population density of each stationary distribution decreases. The state distributions also become increasingly spread out with higher noise. The overall effect of these features is that as noise increases when $r = 0.1$, the probability of the population going extinct in each time period also increases.

5.4 Evaluating closed-loop strategies

In this chapter we have found open-loop solutions to our Markov game model. Recall that solutions to a game are open-loop strategies when they depend only on the current stage

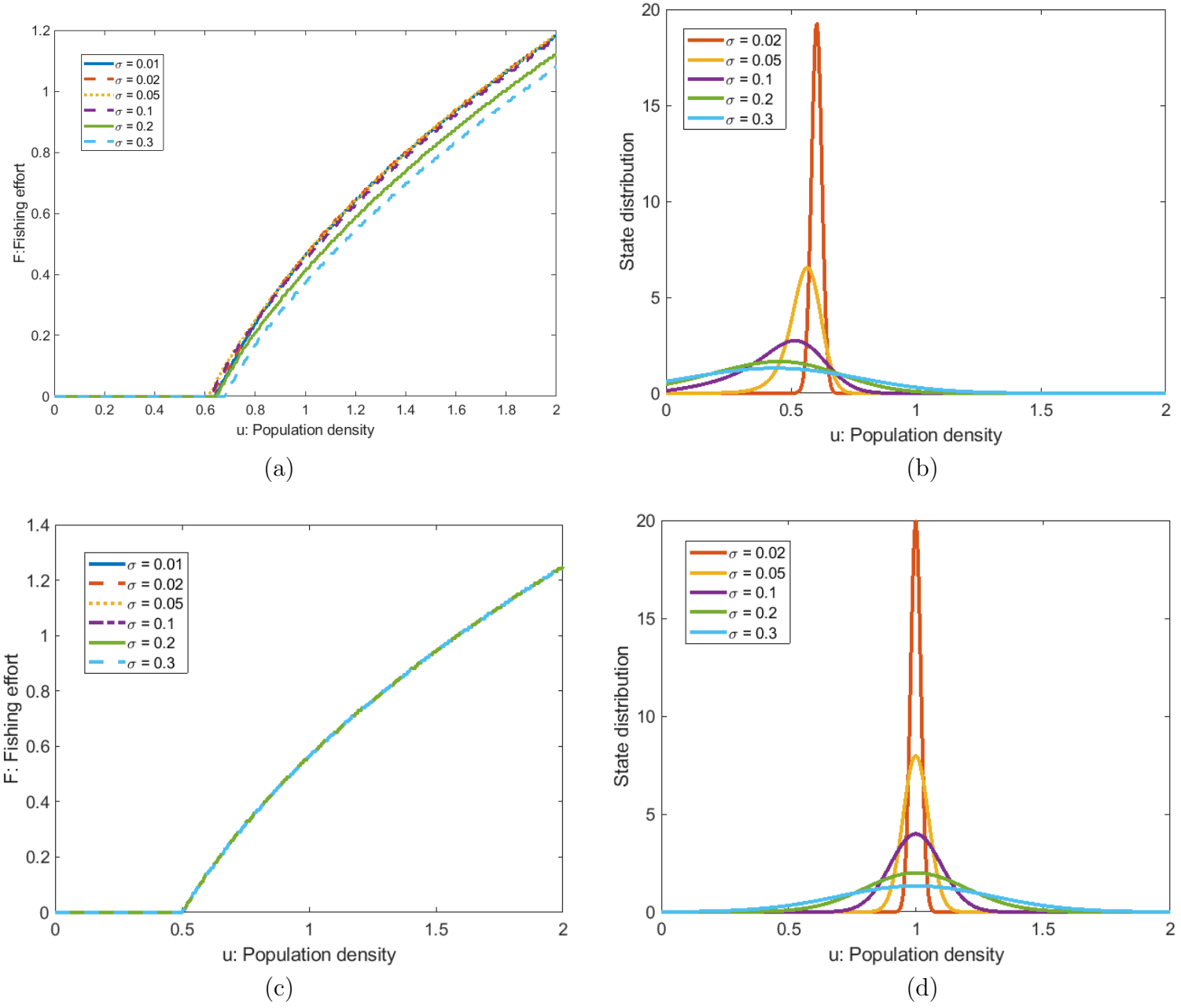


Figure 5.4: (a) and (c) show the stationary Nash equilibrium policies for a range of recruitment noise levels. (b) and (d) show the corresponding stationary state distributions (dominant eigenvalues of M_σ) for those noise levels. In (a) and (b), the population has slow growth with $r = 0.1$. Here the optimal fishing effort at each state slightly decreases with noise. (c) and (d) are for a population with high growth rate $r = 10$, and shows the stochasticity σ in the stock-recruitment relationship does not appear to have a significant effect on optimal policies. Stationary state distributions for the noise level $\sigma = 0.01$ are excluded in both (b) and (d) for graphical reasons. These distributions approach a Dirac-δ function in each case. Other parameter values are $C = 0.2$ and $\delta = 0.9$.

of the iterated game. In our infinite time horizon Markov game, this means that at the beginning of the game a player specifies that they will follow the "optimal" policy to decide which effort they will fish at every year. In contrast, closed-loop strategies allow players to condition their actions at any time on the history of play up to that point. This allows them to "react to exogenous moves by nature, to the realizations of mixed strategies by rivals, and to possible deviations by their rivals from the equilibrium strategies" (Fudenberg and Tirole, 1991). In our scenario, these variations in the history of the game will be expressed as realisations of the noise in the recruitment process. Closed-loop strategies would allow players to choose in any time period an effort level not prescribed by our derived policy, based on their beliefs of their opponent's own strategy. This could enable players to "take advantage" of situations where they believe their rival is playing a certain way, in order to increase their own profits, or alternatively to punish their opponent for acting in a certain way.

The space of potential closed-loop strategies is much larger than the space of open-loop strategies (Fudenberg and Tirole, 1991). Therefore is there is more room for complex strategic behaviour in closed-loop equilibria, as strategies may be dependent on the entire history of play. Closed-loop equilibria tend to be less tractable than open-loop equilibria (Fudenberg and Tirole, 1991). However, we can use the open-loop equilibria we have calculated, along with knowledge of successful strategies in other games, to develop some basic heuristic closed-loop strategies to be investigated. Note that these closed-loop strategies will not in all likelihood be equilibria of our game. However, we can still look at the outcomes of players who use such strategies, to see how the expected payoffs differ from our open-loop optimal policies.

Our development of closed-loop strategies is motivated by results from the iterated prisoner's dilemma (IPD) as discussed in Chapter 3. In an IPD, players can choose to either cooperate or defect in each stage-game. In this game, the Nash equilibrium is to defect in each round, but this will not necessarily maximise the total expected discounted profit over

a number of rounds. Recall that tit-for-tat (TFT) is a strategy for a multi-stage, two-player symmetric game that depends only on the strategies chosen by each player in the previous stage-game. In the first round of an IPD, a player using a TFT strategy will cooperate. After this, they will choose whichever action their opponent used in the previous round. TFT is an effective strategy when in direct competition with many common strategies (Axelrod, 980a). Specifically, when played against itself, it obtains the IPD reward payoff in each stage-game, and when it plays the always defect strategy, it is not defeated, matching it with the punishment strategy (Imhof et al., 2005).

In our Markov game model, we have found two policies that are conceptually similar to the two moves available in a Prisoner's Dilemma. Fishing with a level of effort prescribed by the Nash equilibrium policy (Π^D) we have found in this chapter corresponds to the choice to defect in an IPD. In each case, these choices are the best response to the player's expectations of an opponents behaviour. Fishing at a rate where both player's efforts sum to the Monopolist effort (as calculated using a Markov decision process) is a situation in which the players' actions are on a Pareto frontier. Neither player can unilaterally deviate from such an arrangement in order to increase their own profit, without reducing the other player's profit. Consequently, if each of the two players fishes at half the monopolist effort, the game is Pareto efficient. We refer to this policy as Π^C. Analogously to the method used in Chapter 3, fishing at half the monopolist effort is equivalent to the choice to cooperate in an IPD.

We can use these two policies to define closed-loop strategies, and evaluate them in terms of payoffs and population biomass levels when compared to following each open-loop strategy. To do so, we take inspiration from the strategy TFT (Axelrod, 1987). We define a closed-loop version of TFT for our Markov game as follows:

1. In the first time period, cooperate (i.e. fish with the action $F = \Pi^C(u_1)$)
2. In each successive time period, if you believe the other player cooperated in the previous round of the game, then cooperate. Otherwise defect (by fishing with the action

$F = \Pi^D(u_n)$)

It makes sense that we can similarly define closed-loop strategies which are analogous to other simple strategies for the IPD (which we have discussed in Chapter 3). For instance, an always cooperate ($ALLC$) closed-loop strategy can be defined as

- In each fishing period, fish with effort $F = \Pi^C(u)$.

A closed-loop version of always defect can be defined in a similar manner. It is also clear that we could describe variations of the TFT strategy, such as suspicious tit-for-tat or generous tit-for-tat.

Our closed-loop version of TFT can be applied to a slightly different game than that defined earlier in this chapter. In particular, there are differences in the assumptions about the information available to each player. In the game as defined in Subsection 5.3.2, players have perfect and complete information. At each point in time, each player can observe the value of the state variable u by measuring the population levels. As they know the population biomass at the start of a fishing period (implying that they know exactly the noise in the stock-recruitment process in each year) and they know the yield they obtain during it, they can indirectly measure the fishing effort chosen by the other player. As they know each move of their opponent, the fisher has perfect information. If a fisher used the closed-loop version of TFT in this game, they would be fully aware of whether their actual profits in realisations of the game are due to environmental stochasticity, or due to the fishing efforts of their opponents. A major motivation of this chapter (and this thesis as a whole) is to investigate how uncertainty affects the optimal behaviour of fishers, and what this means for both profits and fish populations. Consequently, we want to adapt our existing game to a situation in which they cannot observe their opponent's behaviour; they will therefore be unable to tell if current population levels are due to noise in recruitment, or exploitation.

We define a slightly different game to evaluate our closed-loop TFT strategy, in order to

satisfy our information requirements. Where players need to choose an action based on whether they believe their opponent has cooperated or defected, they must have uncertainty about whether deviations from their expected payoffs are due to realisations of environmental noise, or due to their opponent playing an unknown strategy. In order to maintain this uncertainty, each player must therefore have noise in their payoff function. We implement this by assuming that players can observe stock levels at the end of each fishing period. With their knowledge of population dynamics, they will therefore be able to determine the expected population density at the beginning of the next fishing period. However, they do not know what the discrepancy from this expected value is at any time, due to noise in recruitment. In this imperfect information game, players use the expected population biomass as the input to their fishing policy. This means that the actual action chosen by the fisher will not necessarily be the optimal action at any given time, and the payoff they receive may differ from their expected payoff. All other facets of the game, including players' beliefs and preferences remain this same in this version of the fishing game.

We use simulation to investigate the relative performance of our open-loop and closed-loop strategies:

1. Population densities are initialised to a random biomass, for which the player knows the expected value

2. Using the chosen strategy, a fishing effort F is selected according to the player's beliefs about the current population state, and their beliefs about their opponent's previous behaviour.

3. The player determines their beliefs about what their immediate reward (R^C) will be if their rival fishes with effort F^C, and what it will be (R^D) if their rival fishes with effort F^D. They use this to define a threshold profit $R^* = (R^C + R^D)/2$.

4. The actual profit R^i for each player is calculated, using the actual population density and the actual fishing effort used by each player.

5. If $R^i > R^*$, player i believes that their rival cooperated in this iteration. If $R^i < R^*$, player i believes that their rival defected in this iteration.

6. Population dynamics are calculated, with the players knowing the expected population state in this next iteration.

7. Noise is added to the population biomass, and we return to step 2.

This process is repeated for a large number of iterations for each of our closed-loop strategies: *ALLC*, *ALLD*, and *TFT*. Figure 5.5 compares the way in which each of our defined closed-loop strategies performs for a single player, when recruitment noise ranges from $\sigma = 0.01$ to $\sigma = 0.3$. (a) shows the average profit per iteration. (b) shows the average fishing effort, (c) the average yield, and (d) the percentage of iterations in which the player chooses to cooperate by fishing according to the Pareto policy Π^C. In the low noise case $\sigma = 0.01$, *TFT* performs identically to *ALLC*, gaining the same profit, and cooperating in 100% of iterations. As noise increases to $\sigma = 0.02$, the average profit decreases slightly, with fishing efforts and yield increasing above the corresponding *ALLC* level. Here cooperation occurs in approximately 75% of iterations. Where $\sigma = 0.05$, profits for a *TFT* player decrease further relative to the *ALLC* level. In fact, profits, efforts, and yields for a *TFT* player occur at approximately midway between the *ALLC* and *ALLD* levels. Furthermore, a *TFT* player will cooperate in 50% of iterations. For higher noise levels, the same relationship between strategies holds as that of $\sigma = 0.05$.

5.5 Discussion

A two-player fishing game with discrete-time population dynamics was used throughout this chapter. Initially, we considered a deterministic version of this game, in order to get a "base-line" estimate of optimal fishing behaviour in the absence of noisy dynamics. The solution method we used was not guaranteed to converge, and indeed did not for some parameter combinations in the deterministic game. In particular, convergence to a stationary policy did

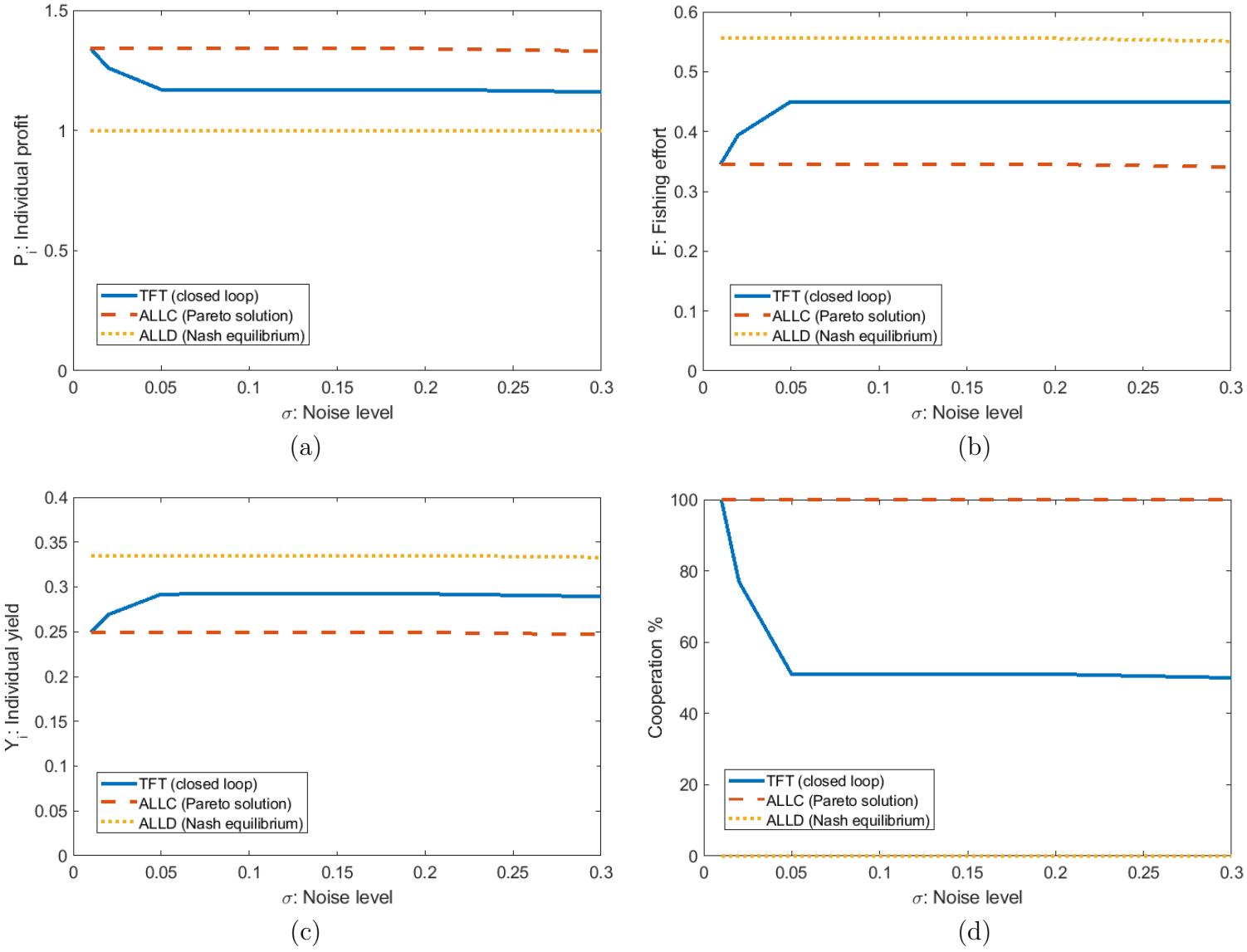


Figure 5.5: The relative performance of the closed loop versions of *TFT*, always cooperate, and always defect are shown here. Results shown are average values per iteration, and are for levels of population noise ranging from $\sigma = 0.01$ to $\sigma = 0.3$. (a) shows the profit obtained by an individual fisher when using each type of closed loop strategy, and is normalised to the value of the Nash equilibrium in the lowest ($\sigma = 0.01$) noise case. (b) shows the fishing efforts used. (c) is the average yield, and (d) shows the percentage of iterations in which cooperation (fishing with the Pareto effort) occurs for each strategy profile. For each result, $r = 10$, $c = 0.5$, and $d = 0.9$.

not occur when the population growth rate was small. We note that two-player games do not necessarily have a deterministic stationary equilibrium (Bowling and Veloso, 2000). However, the symmetric structure of our game, with bounded state and actions spaces, as well as our results from the single player version in Chapter 4, suggest that a stationary deterministic solution may be obtainable, and so we still look for them. In fact, we saw in this chapter that adding a small amount of noise to the game as stochasticity in the stock-recruitment process resulted in convergence regardless of the parameters used. It is therefore likely that the lack of convergence in the deterministic case is a result of numerical noise, potentially caused by the discretisation process. As the number of iterations of the game increased, the effect of this apparent numerical noise on the calculated policies also increased. With any amount of population noise, transition probabilities between states come from a continuous distribution rather than a Dirac-δ function, which we believe dampens the influence of numerical noise on our calculations. It is possible that an alternative solution method, or an adaptation of the current one, could determine a stationary optimal policy in more cases. However, as it stands our model is limited by its non-guaranteed convergence in the deterministic case.

The major purpose of this chapter was to investigate how optimal fishing behaviour in a two-player game changes with noisy dynamics. When the population growth rate was high, we found that the strength of noise had no impact on the Nash equilibrium policies we found. That is, for a given population state, the same level of effort was prescribed by the derived optimal fishing policy, regardless of how noisy the dynamics were. When the population growth rate was small, such that the population had slow dynamics, we found results similar to the monopolist model. That is, as noise increases, the optimal fishing effort at each state decreases. However, this reduction in the optimal fishing effort was in this case small, such that even with slow dynamics there was only a small difference in optimal policies between different noise levels. We associate the reduction in optimal policies with an indirect objective of protecting stock levels into the future such that they can be harvested. As population noise increases, fish stocks become more variable, and so the probability of stocks collapsing

also increases. We believe the introduction of competitive behaviour means that it is no longer worthwhile, in terms of value, for a fisher to try and protect their future earnings in a noisy environment, as their rival can both take advantage of those earnings by increasing their own effort, and also through their actions push the stock towards collapse, so that future earnings are not achievable in any case.

It is clear that regardless of the noise in the stock recruitment relationship, our closed-loop version of *TFT* outperforms the Nash equilibrium (*ALLD*) with respect to the average expected discounted profit obtained by a player using that strategy profile. We developed this strategy heuristically by considering strategies known to be successful in an IPD, and do not claim that it is a closed-loop equilibria of our Markov game. However, we recall that in the case of an IPD, *TFT* is a self-enforcing strategy that is unable to be invaded by *ALLD*. There are many possible closed-loop strategies (Fudenberg and Tirole, 1991), and it may be the case that there are other (non-equilibria) closed-loop strategies that provide better average payoffs than *TFT*.

Chapter 6

Discussion

Fishers make lots of decisions when fishing: when to fish; what gear to use; what size and species of fish to catch, and many more. Throughout this thesis we have used mathematical modelling to investigate how the decisions made by fishers affect the fishery as a whole, and when these decisions lead to cooperative and competitive behaviour. In Chapter 2, we modelled the decisions about which size fish to catch, and whether or not to fish. In Chapters $3-5$ we modelled the decision by fishers about what the best fishing effort is.

In Chapter 2, we modelled a small-scale artisanal fishery in which there are many individual people fishing. Despite the name, small-scale fisheries are responsible for a significant proportion of worldwide employment in fishing sectors, and also in the production of fish for human consumption (FAO, 2014). Descriptions of these types of fisheries vary significantly depend on classification, but generally each individual fisher has a low capacity to fish, has no control over the price they can sell their catch for (or the real value of the fish they and their family eat) (Smith and Basurto, 2019). Fishers in small-scale fisheries make decisions about factors such as the type of gear to use (for example, gillnets or trawls), the time spent fishing, and whether to process the fish they catch (Lopes and Begossi, 2011). While the prevailing image of small-scale fisheries is one of low technology in poor parts of the world (FAO, 2014), small scale fisheries like the conch and lobster fishery in Turks and Caicos have

been able to adapt to technological improvements, resulting in gear changes and increased selectivity over time (Béné and Tewfik, 2001).

The model used in Chapter 2 was built on an existing model (Plank et al., 2016), which was extended to all individual fishers to enter or exit the fishery, based on whether they receive sufficient profits for fishing to be viable. A critical biomass yield was introduced as a proxy for the profit obtained from fishing. In this model, the critical yield is the open-access (bioeconomic equilibrium) solution of the Gordon-Schaefer model (Gordon, 1954). Here there is a large number of fishers, and the total cost of fishing is equal to the total revenue obtained from it. This means that the overall profit is zero accounting for the opportunity costs of fishing. These opportunity costs represent the minimum profit that will keep individuals fishing, rather than convincing them to apply their efforts to other tasks. On average, each fisher receives the minimum profit necessary to continue fishing. In fact, because they obtain equal yields, each fisher in this model obtains equal profit. Empirical evidence suggests that in small-scale fisheries, individual fishers are not always motivated solely by profit, with choices about fishing behaviour being influenced by factors such as community peer pressure, individual fishing skills, and effort reduction (Béné and Tewfik, 2001). However, our assumption in this model that fishers require a minimum profit is reasonable; while individuals may not always act to maximise profit, for them to fish it must still be individually beneficial in terms of satisfying food or income needs.

As in Plank et al. (2016), individual fishers were frequently allowed to change their fishing behaviour. In the model, this was done by allowing each fisher to change their target fish size. In reality, fishers do not have a perfect size-selectivity with respect to the fish they actually catch, and will not have a free and instant ability to change their gear. For instance, individual fishers may own only nets with specific mesh sizes, and may not be able to make quickly new nets to target fish of a different size. On the other hand, aggregate fishing behaviour in small-scale fisheries such as on the Kenyan south coast has shown that diversified

gear selection and fishing behaviours can result in fish being caught from a significant range of the possible sizes of the target species (Tuda et al., 2016). Furthermore, in small-scale fisheries in Brazil, it has been noted that fishers change their behaviour over time based on their observations of the size of other fishers' catches (Lopes and Begossi, 2011). Specifically, when fishers see others with good catches on the previous day, it is suggested they interpret this as an increased possibility of success on the next day. There is a corresponding increase in risk-prone behaviour, which causes an increase in the variability of the yields obtained (Lopes and Begossi, 2011). Our assumption that fishers will change their behaviour for the chance of an increased yield, based on their beliefs about potential yields, therefore seems reasonable.

In our model, it was found that regardless of the starting state of the fishery, the number of active fishers converged over time, to a value dependent on the critical yield. If the critical yield is low, the number of fishers able to be economically sustained by the fishery will be high, while if the required critical yield is high, a small number of fishers will continue to fish. We assumed that there are no barriers to entry in the fishery, such that it has open-access. This assumption may not be perfectly realistic, as there will likely be some costs or time-lags associated with becoming a fisher. However, these costs are likely small compared to the ongoing costs of fishing effort, and indeed if they were included would simply decrease the rate at which individuals join or leave the fishery. As Plank et al. (2016) found similar results for a fixed number of fishers, we do not envision that adding such costs will have a qualitative effect on our results.

The pattern of aggregate size-selectivity at the stable state matched results found in Plank et al. (2016). Fishing pressure was found to be high at body sizes where the productivity of the fish population was high, with the size-selectivity and productivity becoming increasingly proportional over a wide range of body sizes as the total fishing pressure increased. This corresponds to a scenario of balanced harvesting, in which exploitation occurs in pro-

portion to the natural production of the fish population. It is important to note that this balanced harvesting was an emergent outcome of the model; individual fishers self-organised in a manner to maximise their own individual yields, without being explicitly constrained to harvest in a balanced manner.

The biomass spectrum of the exploited fish population had a flattened region over the body sizes which were targeted by fishers. This means that no fisher is able to increase their own mean yield by changing their own size-selectivity behaviour. We conclude that the resultant fishing pattern is an approximate Nash equilibrium. In real terms, the results of this model suggest that in an open-access small-scale fishery, fishers will independently make decisions about their own way to fish (in terms of gear selection, effort, and time), in such a way that they all get a similar catch. Furthermore, without any restrictions or top-down regulatory controls on the fishery, the overall fishing patterns will be such that the there are minimal negative side-effects on the marine ecosystem. This may seem on the surface to be a result that is unlikely to have empirical backing, especially considering that the majority of fisheries worldwide, whether they are small-scale or industrial, are restricted by external controls (Misund et al., 2002). However, there is some evidence which supports the predictions of our model.

In the largely unrestricted fishery based in the Bangweulu swamps in Zambia, fish are harvested over a significant part of their body size range (from 10g to around 10kg), and have been in a stable manner for 50 years (Kolding et al., 2003). Furthermore, in Lake Kariba, (also in Zambia) the overall fishing behaviour in the artisanal fishery is based approximately on natural productivity. Here it has been observed that the fish population has size structure that is close to the size-structure of an unexploited population of the same species, in a different part of the same lake (Kolding and van Zwieten, 2011). Our results of approximate balanced harvesting over a wide range of body sizes therefore seems feasible in an unstructured small-scale fishery, though further investigation and empirical backing will be required.

This model used in Chapter 2 is useful for exploring the emergent outcomes resulting from the aggregate behaviour of a large number of autonomous agents representing a small-scale artisanal fishery. However, the size-structured ecological model that was used is complex, as is the potential information set available to each of the many individual fishers. These factors limit the potential insights the model can give into the possibility for cooperation to arise between fishers. Hence in Chapter 3 we investigated a simpler model combining population ecology and individual fishing behaviour, such that we could focus on insights into the features of cooperation among fishers.

In Chapter 3, we developed a two-player game by well-known unstructured population models with a model of the economic value of fishing. We modelled a situation in which either one or two fishers decide how much fishing effort to expend, for example by making decisions about how many fishing boats they should use, at the beginning of a fishing season. They make these decisions about effort in order to extract as much profit as possible from the fishery. The decision about effort is based only on knowledge about the amount of fishing that the stock can support in total. Understanding the decisions fishers make about their own effort levels is possibly the most important behaviour about fishing to understand (Opaluch and Bockstael, 1984), and is essential for the creation of sustainable fisheries management strategies (Béné and Tewfik, 2001).

The models we used in Chapters $3-5$ all relate to the choices made about fishing efforts in monopolistic (one fisher) and duopolistic (two fisher) scenarios. A monopolistic scenario might represent a situation where there is a single fishing company who has exclusive access to a fish stock, or the entire fishing fleet of a country with strict regulatory controls such that the overall fishing behaviour is highly managed. In a duopoly, there will be two such entities each attempting to harvest the same fish stocks. In either case, the "individual" fishers will in fact be made up of a large number of individual fishing boats and people, with

their aggregate fishing behaviour following a specific pattern. In reality, individual people on fishing boats (whether as part of a monopoly or duopoly) will have their own preferences about how to fish (Béné and Tewfik, 2001), and there may be rulebreaking or deviations from optimal behavior (Watson and Pauly, 2001; Bremner et al., 2009). Whether or not the fishing regulations are followed is a different but related problem that is worth further investigation. This is an issue of cooperative game theory, and as such would require a significantly different modelling process. In this thesis, we have assumed in Chapters $3-5$ that any such regulations or high-level decisions about fishing are perfectly followed at all times, so that idealized overall patterns of exploitation, and their implications, can be investigated.

In the model in Chapter 3, we restricted a continuous choice of possible actions to two possible discrete values. F^C was half the effort at which a monopolist (fisher with sole access) would harvest in order to maximise profit. Two fishers each fishing at this level will result in a socially optimal result, as they equally share high profits, and have a comparatively low level of total exploitation of the fishery. F^D was the level of effort that maximises a player's profit in the next fishing period, with respect to their rival's action. This action is better in terms of the profit immediately gained by the fisher than all other possible effort choices, based on the fisher's beliefs about their competitor. A Nash equilibrium occurs when both fishers use this behaviour.

At equilibrium, it was found that the symmetric two-player game with these two actions was structurally equivalent to a Prisoners Dilemma. This result was resilient to the choice of population model. As such, there is a Nash equilibrium, where both fishers "defect" by fishing with F^D, and a Pareto efficient outcome, where both players "cooperate" by fishing with F^C. It was found that the Pareto outcome not only resulted in greater payoffs for each player, but also had a lower total fishing pressure, and as such a higher equilibrium population density. This means that from an ecological as well as an economic point of view, the Pareto result is preferable to the Nash equilibrium. This result for the two types of solutions

is similar to that found by Levhari and Mirman (1980) in their model.

The model up to this point focused on which decisions to make in a single fishing period. We then investigated a repeated version of the game, in which decisions about fishing effort are made by each player at the beginning of each fishing period. This iterated game was found to be an Iterated Prisoner's Dilemma (IPD) (Tadelis, 2013). Decisions about which effort to choose in subsequent years may be based on the amount of fish caught in previous years. This assumption about changing fishing behaviour based on the past seems reasonable, as similar behaviour has been observed in both small-scale (Lopes and Begossi, 2011) and industrial fisheries, including the New Zealand hoki fishery (Vignaux, 1996). While fishers cannot directly observe their competitors, and do not have perfect information about them, they can measure the amount of fish they personally catch. Their own yield, combined with their beliefs about the way in which their opponent chooses to fish, allows a player the opportunity to (indirectly) respond to the actions of the other player.

Each fisher can choose strategies, or rules that will select their fishing efforts in the future, based on their beliefs about the actions of the other player. The relative profits from choosing various strategies (including $ALLC$, $ALLD$, and TFT) in the IPD were evaluated. When fish populations are able to respond quickly to changes in fishing efforts, the results of these well-known strategies correspond to their known performance in an IPD (Axelrod, 980a; Imhof et al., 2005).

Simulations were used to investigate the game for populations in which growth happens on a slower timescale than players can change their fishing behaviour. With slower dynamics, the relative performance of different strategy profiles changes. For example, with fast dynamics, $ALLD$ v TFT will result in continual defection by the TFT player, leading to the Nash equilibrium, and suboptimal profits for each fisher. With slow dynamics the $ALLD$ player can exploit cooperation by the TFT player, resulting in higher relative payoffs. This

effect occurs because players' have imperfect information in this slow dynamics model, and therefore cannot distinguish between variability in payoff due to population dynamics and due to the actions of the other player.

Players' beliefs about their rival's actions are governed by whether they obtain a yield above a threshold level. The threshold Y^* in this chapter differs in its economic implications from that used in Chapter 2. It does not directly correspond to an economic equilibrium level from the Gordon-Schaefer model. Instead, it represents fishers' expectations. They have uncertainty about actions due to a lack of information, rather than as explicit environmental stochasticity as in Chapter 5. We note that the profit obtained when cooperating is half of the monopolist profit, which is the MEY (maximum economic yield). The profit at the Nash equilibrium when both players compete is greater than the open access solution of the Gordon-Schaefer model, and occurs at a lower effort. The limit of this competitive result when the number of players becomes large is the open access solution, representing a similar type of fishery to that seen in Chapter 2, where anyone is free to exploit the fishery. As Y^* in this chapter is always between the Pareto and Nash yields, the threshold Y^* in this chapter is higher than the corresponding threshold in Chapter 2. In this chapter, we also considered a Beverton-Holt population model, in which growth occurs discretely and instantly at regular intervals, with harvesting happening continuously at a fixed level between these recruitment events. Population dynamics and growth were also separated in this way in Chapters 4 and 5. With Beverton-Holt growth, we again found that the game with restricted efforts was an Iterated Prisoner's Dilemma. This result was invariant to changes in parameter values.

The models used in Chapters 2 and 3 ignore the potential for stochasticity in population dynamics. In reality, noise has a major effect on fish populations (Hjort, 1914; Doherty and Williams, 1988), which can be highly variable in successive years. In Chapters 4 and 5 we therefore considered fish populations with noisy growth, based on the Beverton-Holt population model we considered in Chapter 3.

A stochastic optimisation model was used in Chapter 4 to find how the choice of fishing effort can maximise profit for a monopolist. This was done for the exploitation of a single-species with Beverton-Holt growth dynamics, in which there was noise in the stock-recruitment relationship. The model used was a Markov decision process, which was solved using a dynamic programming method to find optimal fishing policies (Puterman, 2014). These are mappings from each possible level u to the fishing effort F which will maximise the value of the fishery in its current state. When considering the value of the fishery, both the immediate profit obtained from fishing with that effort, and the expected profits that will be available in the future are taken into account. This means that this model allows fishers to consider how their own actions (specifically how they choose to fish) will impact fish stocks in the future. Hence fishers must make decisions about effort which consider the trade-off between catching fish now, and leaving them to reproduce so that they can catch more in the future. To do so, they take into account factors such as the relative value of profit over time, and the likelihood of the population being in a specific state in the future, using their knowledge of how the fish stocks change with time.

We focused on finding stationary optimal policies, which do not change with time, and so are solutions to an infinite time horizon problem (Puterman, 2014). The optimal fishing policies we found for this single-player, unstructured model have some consistent features, regardless of parameter values. Below a certain population biomass, no fishing will occur. This represents a minimum stock level at which a monopolist is willing to fish. When stocks are low, the revenue per unit effort from fishing is small, and so fishing is not profitable. Furthermore, if such a population is further depleted, it is increasingly likely to collapse or go extinct (Jackson et al., 2001). Above this minimum value, the optimal fishing effort increases with population biomass, meaning that the monopolist will choose to fish with greater effort when population levels are high. We also note that stationary policies result in lower rates of exploitation than the profit maximising behaviour in the final year of a finite

time horizon problem . This is because optimal policies take into account the profits that can be acquired from the fish stocks in the future, and so they value some preservation of population abundances.

We also considered how the size of population noise changed the optimal fishing policies. It was found that as the noise in the stock-recruitment relationship increased, the optimal rate of fishing in a a given biomass state decreased. With greater noise, there is a higher probability of the fish population going extinct in any single iteration. With a lower optimal fishing effort, the population biomass will be less depleted in each fishing period, and so it is less likely that the recruitment stochasticity will cause the population to die out. Consequently lower fishing rates will occur with higher noise, in order to preserve the fish population so that some value may be extracted from it in the future.

We are able to compare our results to similar approaches which also use MDPs to model fishing by a monopolist. Mendelssohn (1980) used an unstructured Ricker stock-recruitment model of fish population dynamics, in contrast to our approach using Beverton-Holt dynamics. Importantly, Mendelssohn aimed to maximise the yield obtained from the fishery. As such, there were no costs of fishing, and so this approach can be considered to come from an MSY standpoint rather than with a bioeconomic objective. The policies found in Mendelssohn (1980) always aim to fish down to the biomass point which is the minimum required for fishing. Our results also have such a biomass level, below which fishing will not occur. However, our policies do not deplete population this much, instead having the objective of fishing to a stock level above the minimum required for fishing. Mendelssohn (1978) did consider a fish population with Beverton-Holt growth, but in an age-structured case. Here the optimal fishing policy for each age group was obtained separately. Lane (1989) used a variation of an MDP called a partially observable Markov decision process. Unlike our model, in which interseasonal fishing decisions are made between iterations of the stochastic process, Lane considered the decisions made during a single fishing period.

These decisions were made not just about which effort to fish at but also the location where to fish. A discrete number state space and action space was also used. Allowing variable market prices of fish, Lane found that unit prices have a much more significant impact on the decisions made by fishers than population abundances do. Lane also investigated the effect of improved capacity for stock assessment; it was found that as assessments of stock levels improve, the profits fishers receive also improve (Lane, 1989).

Chapter 5 used Markov games to model the exploitation of a fishery with stochastic dynamics by multiple players. The model was set up in a similar manner to the monopolist Markov model in Chapter 4, but allowed for harvesting by two players with equal access, like in Chapter 3. As before, each player in the duopoly model acts by choosing a fishing effort. Optimal fishing policies were, as in Chapter 4, found by iterating backwards through time. However, within each fishing period there was also iteration between the behaviour chosen by each player, with best responses to the other's strategy successively calculated until players' policies converged within the time period.

For certain parameter combinations, convergence to a stationary policy did not occur for the deterministic two-player version of this model. This lack of convergence was likely due to the impact of numerical noise on the decision problem. For parameter combinations in which convergence did occur, the calculated optimal policies were qualitatively similar to the monopolist version found in Chapter 4, in that there was a minimum biomass for fishing to happen, and that the rate of fishing effort increased with the population level. With stochastic dynamics, there was convergence to an optimal policy for all parameter combinations, suggesting that these probabilistic transitions had a "smoothing" effect on the numerical noise previously preventing convergence.

When the fish population had fast dynamics, it was found that the strength of the noise in the stock-recruitment relationship had no effect on the optimal policy; for any variance

value (with other parameters fixed) the same stationary policy was found. When growth dynamics were slow, we saw results that qualitatively agreed with those from the monopolist model in Chapter 4. The noise level had a direct effect on optimal fishing behaviour, with more noise in the recruitment relationship corresponding to lower optimal fishing efforts at each level of population biomass. However, the strength of this relationship was weaker than in the monopolist model, with there only being a slight decrease in the optimal fishing effort as noise increased from a low level to a high one.

We put forward a possible hypothesis for the difference between the monopolist and two-player results; competition between rival fishers mitigates the ecologically protective behaviour seen with a monopolist. The potential gains, in terms of expected higher future stock levels than can be harvested in the future, will not with certainty go to a fisher who decides to fish less heavily in the immediate term. Instead, their rival could simply fish more in the present term in order to take advantage of the instantaneous profits. With our model formulation, fishers value both guaranteed profits now, and the chance for possible profits in the future. An outcome of the model is that it appears they value the immediate return more in the two-player case.

Fish populations with fast dynamics are replenished quickly after exploitation. This means that high fishing efforts are feasible, as long as the costs incurred are not too great. With heavy exploitation, as long as the fish stock is not exploited to the point of collapse, strong recruitment will mean that stock levels will quickly recover. However, it has been seen that fast growing fish populations are the most at risk of ecological collapse (Pinsky and Byler, 2015). Further investigation will be required, potentially by considering variable or size-specific growth rates. It also be worth investigating how growth rates affect the relative size of MEY, MSY, and unsustainable yields which will lead to population collapse, in more detailed population models.

The optimal policies we found in Chapter 5 were open-loop solutions of the Markov game. Open-loop solutions of the game in which only the state of the fishery at the current point in time are taken into account (Fudenberg and Tirole, 1991). For example, this might mean that at the beginning of the entire fishing process, a player decides that they will follow a specific pattern of fishing effort based on the stock level at a given time. This pattern will not change over time, regardless of how the other player chooses to fish. These differ from closed-loop strategies, with which players are able to take into account the entire history of fishing (Fudenberg and Tirole, 1991), and so can change their behaviour at any time in response to the fishing effort used by the other player. For example, in any fishing period they might fish in a certain way if they believe their opponent had a high effort during the previous period, and fish in a completely different way if they believe their opponent fished with low effort previously.

Influenced by the strategies used in the IPD in Chapter 3, we also developed some closed-loop strategies for this Markov game. Note that these closed-loop strategies are not equilibria of the game; instead we wish to evaluate how more complicated strategies, motivated by existing game theory knowledge, compare to the open-loop solutions in terms of profits and population levels. The main strategy of interest was a closed-loop version of *TFT*. As in Chapter 3, we used a threshold yield to determine players' beliefs about whether their rival cooperated or defected in the previous stage-game. In economic terms, this is the same as in Chapter 3, being midway between the yield obtained at MEY and the yield from the Nash equilibrium. The difference in this chapter is that the players' expectations of yields are not just based on population equilibria; instead they vary over iterations with the current state of the population.

The reason for players' uncertainty about their rivals actions also differs between these chapters. In Chapter 3, this was because of a lack of information about population dynamics. In Chapter 5, it is because of being unaware of the actual realizations of environmental noise.

We compared the average payoff in each iteration when using TFT to the profits when using both the Pareto (half the monopolist) policy, and the Nash equilibrium (open-loop equilibrium) policy. It was found that for low levels of noise, tit-for-tat performed identically to the Pareto policy, while for higher levels of noise tit-for-tat had a lower average payoff than the Pareto policy, but a higher average payoff than the Nash policy. Consequently our closed loop tit-for-tat strategy resulted in higher average profits (at all noise levels) than the open-loop equilibrium.

The assumptions we have made about our Markov game can be compared to assumptions made in similar models, with the types of solutions obtained also able to be compared. Martin-Herran and Rincón-Zapatero (2005) looked for solutions of a continuous-time fishery game that are both Nash equilibria and Pareto efficient. In contrast, we developed Pareto efficient solutions to our game using a monopolist model in Chapter 4, and find Nash equilibria separately in our two-player model. We were however able to compare our Pareto efficient policies to our Nash equilibria, and find that they are typically not the same. Consequently, our results differed in form from those of Martin-Herran and Rincón-Zapatero (2005). A sequential game, in which players take turns to make decisions about fishing at different times was used by Martin-Herran and Rincón-Zapatero (2005). Here, an economic model similar to ours was used, while the information sets of the fishers are asymmetric. Consequently Martin-Herran and Rincón-Zapatero (2005) found different optimal policies for different fishers, unlike our symmetric results. Optimal control methods are used to solve this problem.

The models we have used in Chapters $2-5$ were for a single fish species. Real marine communities have many interacting species, with both conspecific and interspecific predation occurring. Multi-species size spectrum models have been developed (Law et al., 2016), and could be coupled with an individual-based fishing model in which both different target sizes and target species are allowed. Our current fishing model has a fixed fishing effort for

each individual fisher, which may not be a realistic assumption. However, by allowing the number of fishers to change, the total fishing pressure is still variable, and so we would expect that this assumption does not have a significant effect on the aggregate fishing behaviour. We also assumed that there are no barriers to entry or exit of the fishery (whether costs, or lead times.) This is unlikely, but additional costs of this type would simply reduce the rate at which the number of fishers changes. As we have seen in earlier results that with a fixed number of fishers the emergent balanced harvesting outcome still happens this assumption seems reasonable.

In Chapter 3, the fast dynamics case was based on the assumption that the fish population would quickly reach equilibrium under a fixed level of fishing effort, relative to the rate at which fishers are able to change their fishing effort. In reality, many fish species, including some which are of high economic value such as orange roughy (*Hoplostethus atlanticus*), have slow growth rates (Clark, 2001). It is therefore unlikely that fish stocks of these low productivity species will indeed reach equilibrium quickly, especially when compared to the time scales on which fishers would be able to alter their gear or fishing behaviours. The slow dynamics case possibly exemplifies a more realistic scenario for low productivity fish species. In our simulations of this model, decisions about fishing efforts were made based on the yield obtained by each fisher (as a proxy for profit). The yield obtained in each time period was used as an indicator of whether the rival player had cooperated or defected. This allowed the use of well known game-theoretic strategies such as *TFT* (Axelrod, 1987).

The beliefs about which yields should be obtained in specific situations were based on the corresponding fast dynamics equilibria. With slower dynamics, populations will replenish less quickly, and so yields (and profits) should be lower than with fast dynamics. In this chapter, we were aiming to gain understanding of how a fisher's uncertainty about their environment, in terms of population dynamics and other fishers' behaviour, could alter their decisions about fishing. Consequently, while the specific beliefs the individuals have about

expected profits are not necessarily correct in the slow dynamics case, this is representative of fishers not having perfect information about their environment.

In our two-player prisoner's dilemma models, we restricted the actions of each player to a choice between two effort levels. These levels represented complete cooperation, where the maximum economic yield (i.e. the Monopolist profit) is shared equally between the players, and complete competition (the Nash equilibrium), in which each individual fisher chooses the effort that will maximise their profit with respect to the other fisher's effort. The Nash equilibrium effort strictly dominates all other choices of effort in a single-shot prisoner' dilemma, in that whatever action Player i chooses, Player j can increase their own payoff by moving effort towards the Nash level. In reality, a fisher is able to aim for any number of levels of fishing effort, as the combination of a number of inputs include the time spent fishing, the amount of boat-trips taken, and the number of nets used.

Because of the uncertainty in fish stock estimates, parameter values representing populations are unlikely to be perfectly known, and indeed population models do not perfectly represent real fish populations. As such, fishers may not exactly know the fishing effort that leads to the Pareto efficient and Nash equilibria. Moreover, these idealised results are based on equilibrium assumptions that may not hold up to scrutiny in real life fisheries (Larkin, 1977). However, it is reasonable to assume that fishers are aware there is a level of effort where if all fishers cooperate and fish at, population abundances and profits will be high, but that there is a higher level of effort which in the short term will result in increased profits to the fisher who takes advantage of the increased yield, if no other fishers deviate from the cooperative level. This is the essential information that is captured in our restriction to two actions; the outcomes themselves do not need to be optimal in terms of maximising available profit, as long as this trade-off between cooperating, and defecting for short-term profit is captured by the model.

We have also assumed that the two fishers are perfectly symmetric, with identical costs of fishing, and identical preferences and information. If that was not the case, the levels of effort for each solution to the prisoner's dilemma games would not be symmetric, and so different fishers could exhibit different behaviour in certain situations. For example, fishers may not have the same expectations of profit, and so there may be different levels at which slow population dynamics cause one fisher to believe that defection from the Pareto frontier has occurred. Furthermore, we have assumed that all fish are perfectly and equally catchable, with there being no spatial variation in fish population density, and no effect of population depletion on the ability to catch fish. Setting a "fixed" level of fishing effort will not in actuality translate to a guaranteed yield for an individual fisher.

In the stochastic population model used in Chapter 4, all noise was incorporated discretely into the stock-recruitment relationship. We could have included an additional continuous source of noise, affecting the population dynamics during harvesting. This would be able to represent sources of stochasticity with temporal effects, such as seasonal variations in population catchability or fishing patterns. Furthermore, the existing noise in recruitment is density-independent. It is possible that the actual population level has an effect not only on the expected population size in the next time period, but also in the variability in recruitment. A multiplicative noise term in the population model would be able to represent this type of variation. The MDP used to obtain optimal fishing policies for the stochastic population was solved under the assumption of an infinite time horizon, with constant parameter values. This implies that the population does not characteristically change between time periods of the stochastic process; there are no long-term impacts of exploitation or any time dependence on population dynamics.

It is also possible that fishing regimes could change over time. This could include new entrants to the fishery, in which there is no longer sole access to the fish stock. In this case, our Markov game model would be appropriate to use. In fact, the results from the

monopolist model are still used to determine the Pareto frontier in the Markov game model, and so are still of use when applied to fisheries in which there are shared stocks. It is also possible that there are changes in the costs of fishing or the price of fish over time. This could be included in the stochastic model by allowing time-dependent parameter values. For the current formulation of our MDP model, with bounded state and action spaces, a solution (a stationary optimal policy) is guaranteed to exist, but is not guaranteed to be unique. There could therefore be other policies which are of the same value to the sole fisher. On the other hand, Markov games will not necessarily have a stationary solution.

The Markov games in Chapter 5 were solved by finding open-loop equilibria rather than closed-loop. The space of closed-loop strategies is much larger, but finding closed-loop equilibria is comparatively difficult (Puterman, 2014). We were however able to define a simple closed-loop strategy based on TFT that performed better than the open-loop equilibrium that we did calculate. While not derived as a profit-maximising equilibrium strategy, this closed-loop TFT resulted in a higher average profit per fishing period than the open-loop Nash solution, for every level of noise in the population dynamics. This result suggests that further investigation of closed-loop strategies could be worthwhile for this problem, and that the use of strategies motivated by results from Chapter 3 is a reasonable starting point. Adaptations of TFT, such as $STFT$ and $GTFT$ may be considered. A potentially worthwhile extension to this model would be to allow mixed strategies (which includes the aforementioned $GTFT$), where the optimal action in each state is taken from a probability distribution, rather than from a deterministic mapping. A possible alternative approach would be to look for cyclic equilibria; these solutions are not stationary, but satisfy the condition for an equilibrium in that there is no benefit for a player to unilaterally deviate from it (Zinkevich et al., 2006). Cyclic equilibria are a set of stationary policies which are strictly cycled through (Zinkevich et al., 2006). Another valuable extension would be to include some size-structure in the population model. While this would not necessarily need to be a full size-spectrum model like in Chapter 2, allowing fishers to target individual fish of dif-

ferent sizes at different rates would add some biological complexity to the model, and would possibly allow an investigation of competing fishers using balanced harvesting strategies in a noisy environment.

Bibliography

Allen, D. M. and Barker, D. L. (1990). Interannual variations in larval fish recruitment to estuarine epibenthic habitats. *Marine Ecology Progress Series. Oldendorf*, 63(2):113–125.

Allison, E. H. and Ellis, F. (2001). The livelihoods approach and management of small-scale fisheries. *Marine Policy*, 25(5):377–388.

Andersen, K. H., Jacobsen, N. S., and Farnsworth, K. D. (2015). The theoretical foundations for size spectrum models of fish communities. *Canadian Journal of Fisheries and Aquatic Sciences*, 73(4):575–588.

Axelrod, R. (1980a). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24(1):3–25.

Axelrod, R. (1980b). More effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24(3):379–403.

Axelrod, R. (1981). The emergence of cooperation among egoists. *American Political Science Review*, 75(2):306–318.

Axelrod, R. (1987). The evolution of strategies in the iterated prisoners dilemma. *The Dynamics of Norms*, pages 1–16.

Bailey, M., Sumaila, U. R., and Lindroos, M. (2010). Application of game theory to fisheries over three decades. *Fisheries Research*, 102(1):1–8.

Baranov, F. I. (1918). *On the Question of the Biological Basis of Fisheries*. Bloomington Ind., Indiana University.

Baum, J. K., Myers, R. A., Kehler, D. G., Worm, B., Harley, S. J., and Doherty, P. A. (2003). Collapse and conservation of shark populations in the Northwest Atlantic. *Science*, 299(5605):389–392.

Beddington, J. R., Agnew, D. J., and Clark, C. W. (2007). Current problems in the management of marine fisheries. *Science*, 316(5832):1713–1716.

Bellman, R. (1966). Dynamic programming. *Science*, 153(3731):34–37.

Bellman, R. et al. (1954). The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 60(6):503–515.

Béné, C. and Tewfik, A. (2001). Fishing effort allocation and fishermen's decision making process in a multi-species small-scale fishery: analysis of the conch and lobster fishery in Turks and Caicos Islands. *Human Ecology*, 29(2):157–186.

Berezansky, L. and Braverman, E. (2004). On impulsive Beverton-Holt difference equations and their applications. *Journal of Difference Equations and Applications*, 10(9):851–868.

Berkes, F. (1985). Fishermen and the tragedy of the commons. *Environmental Conservation*, 12(3):199–206.

Beveridge, M. C., Thilsted, S., Phillips, M., Metian, M., Troell, M., and Hall, S. (2013). Meeting the food and nutrition needs of the poor: the role of fish and the opportunities and challenges emerging from the rise of aquaculture. *Journal of Fish Biology*, 83(4):1067–1084.

Beverton, R. J. and Holt, S. J. (1957/2012). *On the dynamics of exploited fish populations*, volume 11. Springer Science & Business Media.

Birkeland, C. and Dayton, P. K. (2005). The importance in fishery management of leaving the big ones. *Trends in Ecology & Evolution*, 20(7):356–358.

Booth, D. and Beretta, G. (1994). Seasonal recruitment, habitat associations and survival of pomacentrid reef fish in the US Virgin Islands. *Coral Reefs*, 13(2):81–89.

Borrell, B. (2013). A big fight over little fish. *Nature*, 493(7434):597–598.

Botsford, L. W., Castilla, J. C., and Peterson, C. H. (1997). The management of fisheries and marine ecosystems. *Science*, 277(5325):509–515.

Bowling, M. and Veloso, M. (2000). An analysis of stochastic game theory for multiagent reinforcement learning. Technical report, Carnegie-Mellon Univ, Pittsburgh P.A., School of Computer Science.

Boyd, R. and Lorberbaum, J. P. (1987). No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature*, 327(6117):58–59.

Brauer, F., Castillo-Chavez, C., and Castillo-Chavez, C. (2001). *Mathematical Models in Population Biology and Epidemiology*, volume 40. Springer.

Bremner, G., Johnstone, P., Bateson, T., and Clarke, P. (2009). Unreported bycatch in the New Zealand west coast South Island hoki fishery. *Marine Policy*, 33(3):504–512.

Brewer, D., Heales, D., Milton, D., Dell, Q., Fry, G., Venables, B., and Jones, P. (2006). The impact of turtle excluder devices and bycatch reduction devices on diverse tropical marine communities in Australia's northern prawn trawl fishery. *Fisheries Research*, 81(2):176–188.

Burgess, M. G., Diekert, F. K., Jacobsen, N. S., Andersen, K. H., and Gaines, S. D. (2015). Remaining questions in the case for balanced harvesting. *Fish and Fisheries*.

Campana, S. E. (1996). Year-class strength and growth rate in young Atlantic cod Gadus morhua. *Marine Ecology Progress Series*, 135:21–26.

Chiarella, C., Kemp, M. C., Van Long, N., and Okuguchi, K. (1984). On the economics of international fisheries. In *Contributions to Economic Analysis*, volume 150, pages 189–198. Elsevier.

Clark, M. (2001). Are deepwater fisheries sustainable? the example of orange roughy (Hoplostethus atlanticus) in New Zealand. *Fisheries Research*, 51(2-3):123–135.

Conover, D. O. and Present, T. M. (1990). Countergradient variation in growth rate: compensation for length of the growing season among atlantic silversides from different latitudes. *Oecologia*, 83(3):316–324.

Coop, P. and Brunckhorst, D. (1999). Triumph of the commons: age-old participatory practices provide lessons for institutional reform in the rural sector. *Australian Journal of Environmental Management*, 6(2):69–77.

Cox, S. J. B. (1985). No tragedy of the commons. *Environmental Ethics*, 7(1):49–61.

Crowder, L. B., Crouse, D. T., Heppell, S. S., and Martin, T. H. (1994). Predicting the impact of turtle excluder devices on loggerhead sea turtle populations. *Ecological Applications*, 4(3):437–445.

Davies, R., Cripps, S., Nickson, A., and Porter, G. (2009). Defining and estimating global marine fisheries bycatch. *Marine Policy*, 33(4):661–672.

Davis, M. W. and Ryer, C. H. (2003). Understanding fish bycatch discard and escapee mortality. *AFSC Quartely Report, Seattle*, pages 1–9.

Dixit, A. K., Sherrerd, J. J., et al. (1990). *Optimization in Economic Theory*. Oxford University Press on Demand.

Dockner, E., Feichtinger, G., and Mehlmann, A. (1989). Noncooperative solutions for a differential game model of fishery. *Journal of Economic Dynamics and Control*, 13(1):1–20.

Doherty, P. J. and Williams, D. M. (1988). The replenishment of coral reef fish populations. *Oceanogr Mar Biol Annu Rev*, 26(48):551.

FAO (2014). *The State of World Fisheries and Aquaculture, 2014*. Food & Agriculture Organization of the United Nations.

Feeny, D., Hanna, S., and McEvoy, A. F. (1996). Questioning the assumptions of the "tragedy of the commons" model of fisheries. *Land Economics*, pages 187–205.

Filar, J. and Vrieze, K. (2012). *Competitive Markov Decision Processes*. Springer Science & Business Media.

Fogarty, M. J., Sissenwine, M. P., and Cohen, E. B. (1991). Recruitment variability and the dynamics of exploited marine populations. *Trends in Ecology & Evolution*, 6(8):241–246.

Fortibuoni, T., Libralato, S., Arneri, E., Giovanardi, O., Solidoro, C., and Raicevich, S. (2017). Fish and fishery historical data since the 19th century in the Adriatic Sea, Mediterranean. *Scientific Data*, 4:170104.

Froese, R., Walters, C., Pauly, D., Winker, H., Weyl, O. L., Demirel, N., Tsikliras, A. C., and Holt, S. J. (2015). A critique of the balanced harvesting approach to fishing. *ICES Journal of Marine Science: Journal du Conseil*, page fsv122.

Fudenberg, D. and Tirole, J. (1991). *Game Theory*. Cambridge, Massachusetts. MIT press.

Garcia, S., Kolding, J., Rice, J., Rochet, M.-J., Zhou, S., Arimoto, T., Beyer, J., Borges, L., Bundy, A., Dunn, D., et al. (2012). Reconsidering the consequences of selective fisheries. *Science*, 335(6072):1045–1047.

Gordon, H. S. (1954). The economic theory of a common-property resource: the fishery. In *Classic Papers in Natural Resource Economics*, pages 178–203. Springer.

Hardin, G. (1968). The tragedy of the commons. *Journal of Natural Resources Policy Research*, 1(3):243–253.

Hjort, J. (1914). Fluctuations in the great fisheries of northern europe viewed in the light of biological research. In *Rapports et Procs-verbaux des Réunions*. ICES.

Hjort, J. (1933). The optimum catch. *Hvalradets Skrifter*, 7:92–127.

Houde, E. and Hoyt, R. (1987). Fish early life dynamics and recruitment variability. *Trans. Am. Fish. Soc.*

Howard, R. A. (1960). *Dynamic Programming and Markov Processes.* Oxford, John Wiley.

Hu, J. and Wellman, M. P. (2003). Nash Q-learning for general-sum stochastic games. *Journal of machine learning research*, 4(Nov):1039–1069.

Hu, J., Wellman, M. P., et al. (1998). Multiagent reinforcement learning: theoretical framework and an algorithm. In *ICML*, volume 98, pages 242–250. Citeseer.

Huxley, T. (1883). Inaugural address. fisheries exhibition, london. *London: W. Clowes and Sons*.

Imhof, L. A., Fudenberg, D., and Nowak, M. A. (2005). Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences*, 102(31):10797–10800.

Jackson, J. B., Kirby, M. X., Berger, W. H., Bjorndal, K. A., Botsford, L. W., Bourque, B. J., Bradbury, R. H., Cooke, R., Erlandson, J., Estes, J. A., et al. (2001). Historical overfishing and the recent collapse of coastal ecosystems. *Science*, 293(5530):629–637.

Jacobsen, N. S., Gislason, H., and Andersen, K. H. (2014). The consequences of balanced harvesting of fish communities. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1775):20132701.

Jobling, M. (1995). Fish bioenergetics. *Oceanographic Literature Review*, 9(42):785.

Kaitala, V. and Pohjola, M. (1988). Optimal recovery of a shared resource stock: a differential game model with efficient memory equilibria. *Natural Resource Modeling*, 3(1):91–119.

Kawarazuka, N. and Béné, C. (2011). The potential role of small fish species in improving micronutrient deficiencies in developing countries: building evidence. *Public Health Nutrition*, 14(11):1927–1938.

Kolding, J., Ticheler, H., and Chanda, B. (2003). The Bangweulu swamps - a balanced small-scale multispecies fishery. *FAO Fisheries Technical Paper*, 426(2):34–66.

Kolding, J. and van Zwieten, P. A. (2011). The tragedy of our legacy: how do global management discourses affect small scale fisheries in the south? In *Forum for development Studies*, volume 38, pages 267–297. Taylor & Francis.

Kolding, J., van Zwieten, P. A., and Mosepele, K. (2016). Where there is water, there is fish. small-scale inland fisheries in Africa: Dynamics and importance. *A History of Water: Series III, Volume 3: Water and Food*, page 439.

Kot, M. (2001). *Elements of Mathematical Ecology*. Cambridge University Press.

Kraines, D. P. and Kraines, V. Y. (2000). Natural selection of memory-one strategies for the iterated prisoner's dilemma. *Journal of Theoretical Biology*, 203(4):335–355.

Lane, D. E. (1989). A partially observable model of decision making by fishermen. *Operations Research*, 37(2):240–254.

Larkin, P. A. (1977). An epitaph for the concept of maximum sustained yield. *Transactions of the American Fisheries Society*, 106(1):1–11.

Law, R., Kolding, J., and Plank, M. J. (2013). Squaring the circle: reconciling fishing and conservation of aquatic ecosystems. *Fish and Fisheries*.

Law, R., Plank, M. J., and Kolding, J. (2012). On balanced exploitation of marine ecosystems: results from dynamic size spectra. *ICES Journal of Marine Science: Journal du Conseil*, 69(4):602–614.

Law, R., Plank, M. J., and Kolding, J. (2016). Balanced exploitation and coexistence of interacting, size-structured, fish species. *Fish and Fisheries*, 17(2):281–302.

Lehodey, P., Alheit, J., Barange, M., Baumgartner, T., Beaugrand, G., Drinkwater, K., Fromentin, J.-M., Hare, S., Ottersen, G., Perry, R., et al. (2006). Climate variability, fish, and fisheries. *Journal of Climate*, 19(20):5009–5030.

Levhari, D. and Mirman, L. J. (1980). The great fish war: an example using a dynamic Cournot-Nash solution. *The Bell Journal of Economics*, pages 322–334.

Levin, S. A. (1976). Population dynamic models in heterogeneous environments. *Annual Review of Ecology and Systematics*, 7(1):287–310.

Littman, M. L. (1994). Markov games as a framework for multi-agent reinforcement learning. In *Machine Learning Proceedings 1994*, pages 157–163. Elsevier.

Littman, M. L. (2001). Value-function reinforcement learning in Markov games. *Cognitive Systems Research*, 2(1):55–66.

Lomnicki, A. (1978). Individual differences between animals and the natural regulation of their numbers. *The Journal of Animal Ecology*, pages 461–475.

Longley, C., Haraksingh Thilsted, S., Beveridge, M., Cole, S., Banda Nyirenda, D., Heck, S., and Hother, A.-L. (2014). The role of fish in the first 1,000 days in Zambia. *IDO Special Collection*.

Lopes, P. and Begossi, A. (2011). Decision-making processes by small-scale fishermen on the southeast coast of Brazil. *Fisheries Management and Ecology*, 18(5):400–410.

MacLennan, D. N. (1992). Fishing gear selectivity: an overview. *Fisheries Research*, 13(3):201–204.

Malthus, T. R. (1798). *An Essay on the Principle of Population: Or, a View of Its Past and Present Effects on Human Happiness, With an Inquiry Into Our Prospects Respecting*

the Future Removal or Mitigation of the Evils Which it Occasions. London, Reeves and Turner.

Mann, S. H. (1970). A mathematical theory for the harvest of natural animal populations when birth rates are dependent on total population size. *Mathematical Biosciences*, 7(1-2):97–110.

Martin-Herran, G. and Rincón-Zapatero, J. P. (2005). Efficient Markov perfect Nash equilibria: theory and application to dynamic fishery games. *Journal of Economic Dynamics and Control*, 29(6):1073–1096.

McKelvey, R. (1997). Game-theoretic insights into the international management of fisheries. *Natural Resource Modeling*, 10(2):129–171.

Mendelssohn, R. (1978). Optimal harvesting strategies for stochastic single-species, multiage class models. *Mathematical Biosciences*, 41(3-4):159–174.

Mendelssohn, R. (1980). Using Markov decision models and related techniques for purposes other than simple optimization: analyzing the consequences of policy alternatives on the management of salmon runs. *Fishery Bull.*, 78(1):35–50.

Mendelssohn, R. (1982). Discount factors and risk aversion in managing random fish populations. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(9):1252–1257.

Mills, D. J., Westlund, L., de Graaf, G., Kura, Y., Willman, R., and Kelleher, K. (2011). Under-reported and undervalued: small-scale fisheries in the developing world. *Small-scale Fisheries management: Frameworks and Approaches for the Developing World*, pages 1–15.

Misund, O. A., Kolding, J., and Fréon, P. (2002). Fish capture devices in industrial and artisanal fisheries and their influence on management. *Handbook of Fish Biology and Fisheries*, 2:13–36.

Munro, G. R. (1979). The optimal management of transboundary renewable resources. *Canadian Journal of Economics*, pages 355–376.

Myers, R. A. and Worm, B. (2003). Rapid worldwide depletion of predatory fish communities. *Nature*, 423(6937):280.

Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, pages 286–295.

Nash, J. (1953). Two-person cooperative games. *Econometrica: Journal of the Econometric Society*, pages 128–140.

Nowak, M., Sigmund, K., et al. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364(6432):56–58.

O'Brien, C. M., Fox, C. J., Planque, B., and Casey, J. (2000). Fisheries: climate variability and North Sea cod. *Nature*, 404(6774):142.

Opaluch, J. J. and Bockstael, N. E. (1984). Behavioral modeling and fisheries management. *Marine Resource Economics*, 1(1):105–115.

Pawson, M. (1990). Using otolith weight to age fish. *Journal of Fish Biology*, 36(4):521–531.

Pearl, R. and Reed, L. J. (1920). On the rate of growth of the population of the united states since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences of the United States of America*, 6(6):275.

Pepin, P. and Myers, R. A. (1991). Significance of egg and larval size to recruitment variability of temperate marine fish. *Canadian Journal of Fisheries and Aquatic Sciences*, 48(10):1820–1828.

Pinsky, M. L. and Byler, D. (2015). Fishing, fast growth and climate variability increase the risk of collapse. *Proceedings of the Royal Society B: Biological Sciences*, 282(1813):20151053.

Plank, M. J., Kolding, J., Law, R., Gerritsen, H. D., and Reid, D. (2016). Balanced harvesting can emerge from fishing decisions by individual fishers in a small-scale fishery. *Fish and Fisheries*.

Powell, W. B. (2007). *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, volume 703. John Wiley & Sons.

Pritt, J. J., Roseman, E. F., and O'brien, T. P. (2014). Mechanisms driving recruitment variability in fish: comparisons between the Laurentian Great Lakes and marine systems. *ICES Journal of Marine Science*, 71(8):2252–2267.

Puterman, M. L. (2014). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.

Rapoport, A. and Chammah, A. M. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation*, volume 165. University of Michigan press.

Ricker, W. E. (1954). Stock and recruitment. *Journal of the Fisheries Board of Canada*, 11(5):559–623.

San Martin, E., Irigoien, X., Harris, R. P., López-Urrutia, A., Zubkov, M. V., and Heywood, J. L. (2006). Variation in the transfer of energy in marine plankton along a productivity gradient in the Atlantic Ocean. *Limnology and Oceanography*, 51(5):2084–2091.

Schaefer, M. B. (1954). Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Inter-American Tropical Tuna Commission Bulletin*, 1(2):23–56.

Schiesser, W. E. (2012). *The Numerical Method of Lines: Integration of Partial Differential Equations*. Elsevier.

Sethi, S. A., Branch, T. A., and Watson, R. (2010). Global fishery development patterns are driven by profit but not trophic level. *Proceedings of the National Academy of Sciences*, 107(27):12163–12167.

Shapley, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100.

Smith, H. and Basurto, X. (2019). Defining small-scale fisheries and examining the role of science in shaping perceptions of who and what counts: A systematic review. *Frontiers in Marine Science*, 6:236.

Smith, J. M. (1982). *Evolution and the Theory of Games*. Cambridge university press.

Smith, J. M. and Price, G. R. (1973). The logic of animal conflict. *Nature*, 246(5427):15.

Sobel, M. (1982). Stochastic fishery games with myopic equilibria. *Essays in the Economics of Renewable Resources*, pages 259–268.

Sobel, M. J. (1981). Myopic solutions of Markov decision processes and stochastic games. *Operations Research*, 29(5):995–1009.

Sumaila, U. R. (1997). A review of game theoretic models of fishing. *Marine Policy*.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT press.

Swearer, S. E., Caselle, J. E., Lea, D. W., and Warner, R. R. (1999). Larval retention and recruitment in an island population of a coral-reef fish. *Nature*, 402(6763):799.

Tadelis, S. (2013). *Game Theory: An Introduction*. Princeton University Press.

Tsikliras, A. C. and Froese, R. (2019). Maximum sustainable yield. *Encyclopedia of Ecology,*, pages 1–20.

Tsikliras, A. C. and Polymeros, K. (2014). Fish market prices drive overfishing of the big ones. *PeerJ*, 2:e638.

Tuda, P. M., Wolff, M., and Breckwoldt, A. (2016). Size structure and gear selectivity of target species in the multispecies multigear fishery of the kenyan south coast. *Ocean & Coastal Management*, 130:95–106.

Tyler, J. A. and Rose, K. A. (1994). Individual variability and spatial heterogeneity in fish population models. *Reviews in Fish Biology and Fisheries*, 4(1):91–123.

Ursin, E. (1973). *On the Prey Size Preferences of Cod and Dab*. Danmarks Fiskeri-og Havundersøgelser.

Verhulst, P.-F. (1838). Notice sur la loi que la population suit dans son accroissement. *Corresp. Math. Phys.*, 10:113–126.

Vignaux, M. (1996). Analysis of vessel movements and strategies using commercial catch and effort data from the New Zealand hoki fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, 53(9):2126–2136.

Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton university press.

Von Stackelberg, H. (2010). *Market Structure and Equilibrium*. Springer Science & Business Media.

Vrieze, O. J. (1987). Stochastic games with finite state and action spaces. *CWI tracts*.

Watson, R. and Pauly, D. (2001). Systematic distortions in world fisheries catch trends. *Nature*, 414(6863):534.

White, D. J. (1985). Real applications of Markov decision processes. *Interfaces*, 15(6):73–83.

White, D. J. (1993). A survey of applications of Markov decision processes. *Journal of the Operational Research Society*, 44(11):1073–1096.

Wiens, J. A. (1976). Population responses to patchy environments. *Annual Review of Ecology and Systematics*, 7(1):81–120.

Wolff, M., Taylor, M., and Tesfaye, G. (2015). Implications of using small meshed gillnets for the sustainability of fish populations: a theoretical exploration based on three case studies. *Fisheries Management and Ecology*, 22(5):379–387.

Zhou, S., Smith, A. D., Punt, A. E., Richardson, A. J., Gibbs, M., Fulton, E. A., Pascoe, S., Bulman, C., Bayliss, P., and Sainsbury, K. (2010). Ecosystem-based fisheries management requires a change to the selective fishing philosophy. *Proceedings of the National Academy of Sciences*, 107(21):9485–9489.

Zinkevich, M., Greenwald, A., and Littman, M. L. (2006). Cyclic equilibria in Markov games. In *Advances in Neural Information Processing Systems*, pages 1641–1648.